Things you don't want to see in your diffraction data

(Why we do want to see diffraction patterns in 3D)

Andrey Lebedev, CCP4

A questions to ask if structure determination or refinement does not go smooth:

Is there any crystal pathology? (Crystal pathology = no global periodicity)

If yes:

- successful structure solution is less likely
- even if the structure is solved, bad refinement stats are likely
  - explain in the manuscript
    (e.g. a picture from dials.reciprocal_lattice_viewer)
- consider the option of finding better crystallisation conditions

Merohedral twinning cannot be detected by direct examination of diffraction patterns. There are other crystal pathologies that can be.

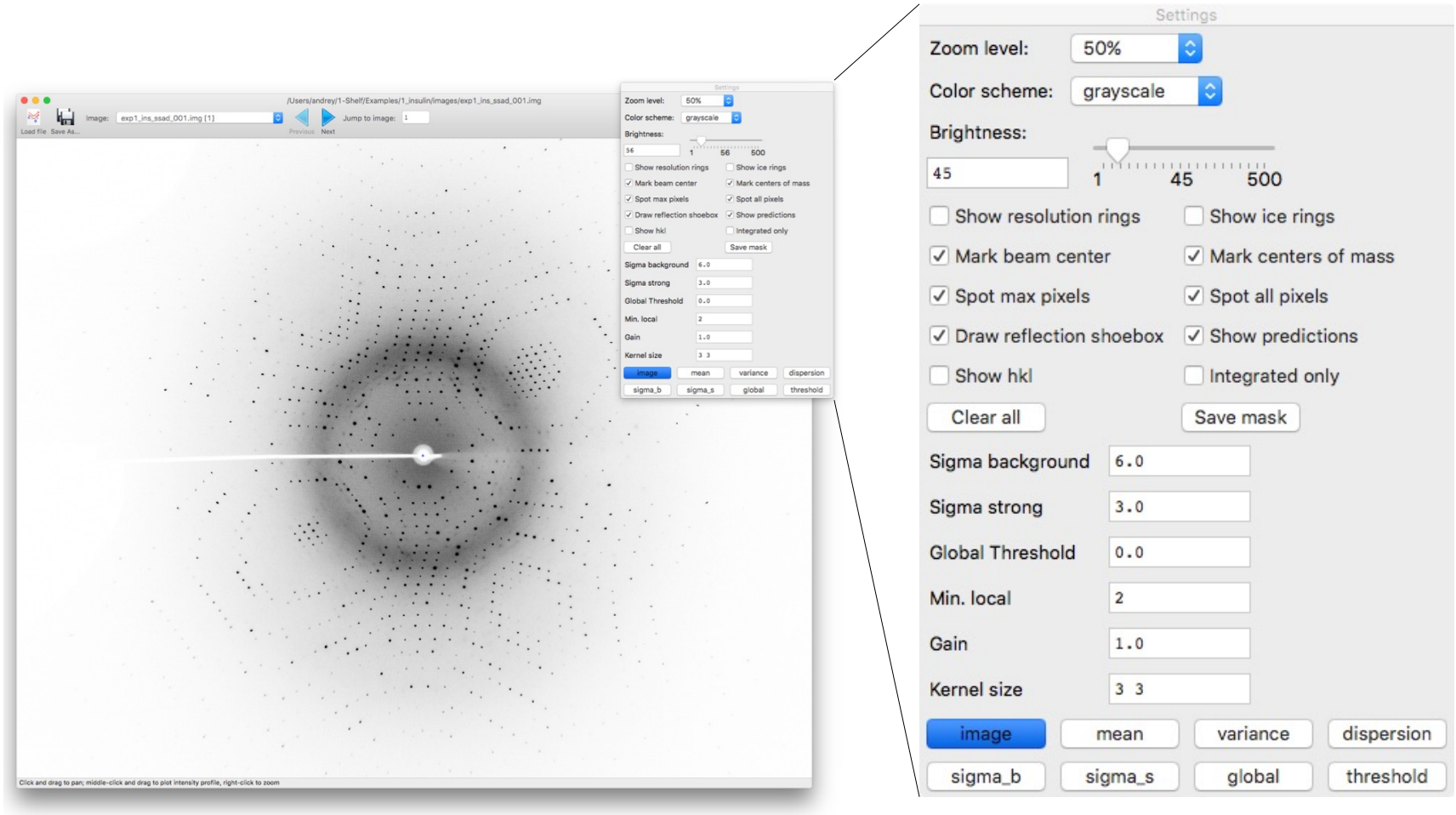
A few examples are presented here.

If my explanation is not clear enough – look at the pictures anyway

- Graphical facilities in DIALS
- Inter-grown crystals
- OD-structures
- Partially disordered OD-structures
- Pseudo-translation

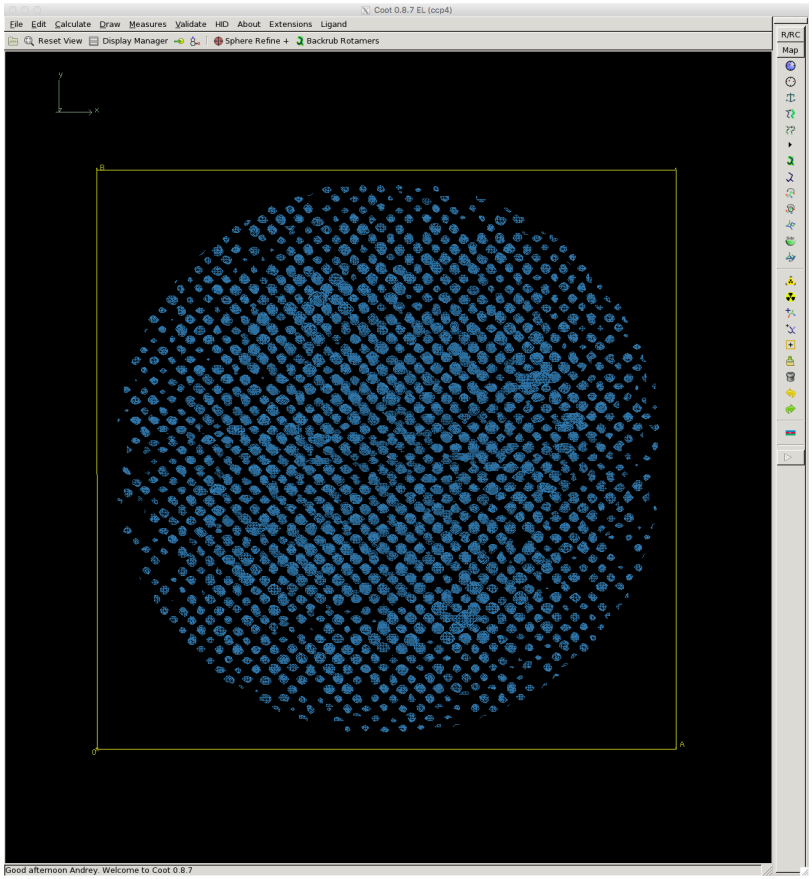
Image viewer

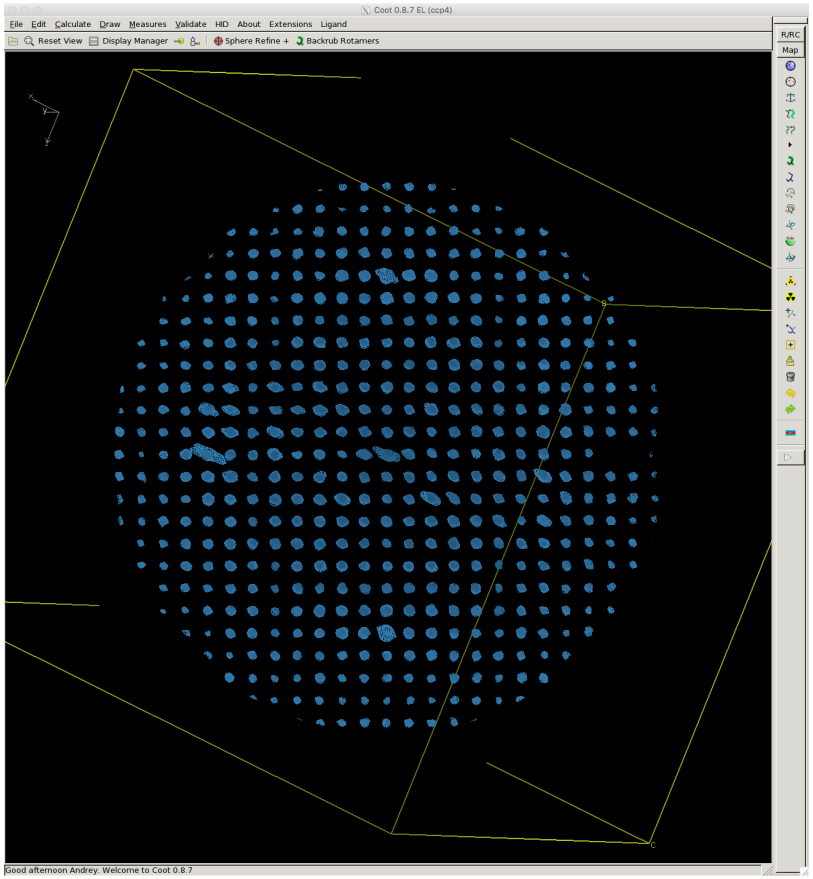
Cubic insulin, the experiment 1 from HZB MX tutorial



Sweep of images as 3D map

```
dials.rs_mapper data_00*.img
coot --map output.ccp4
```

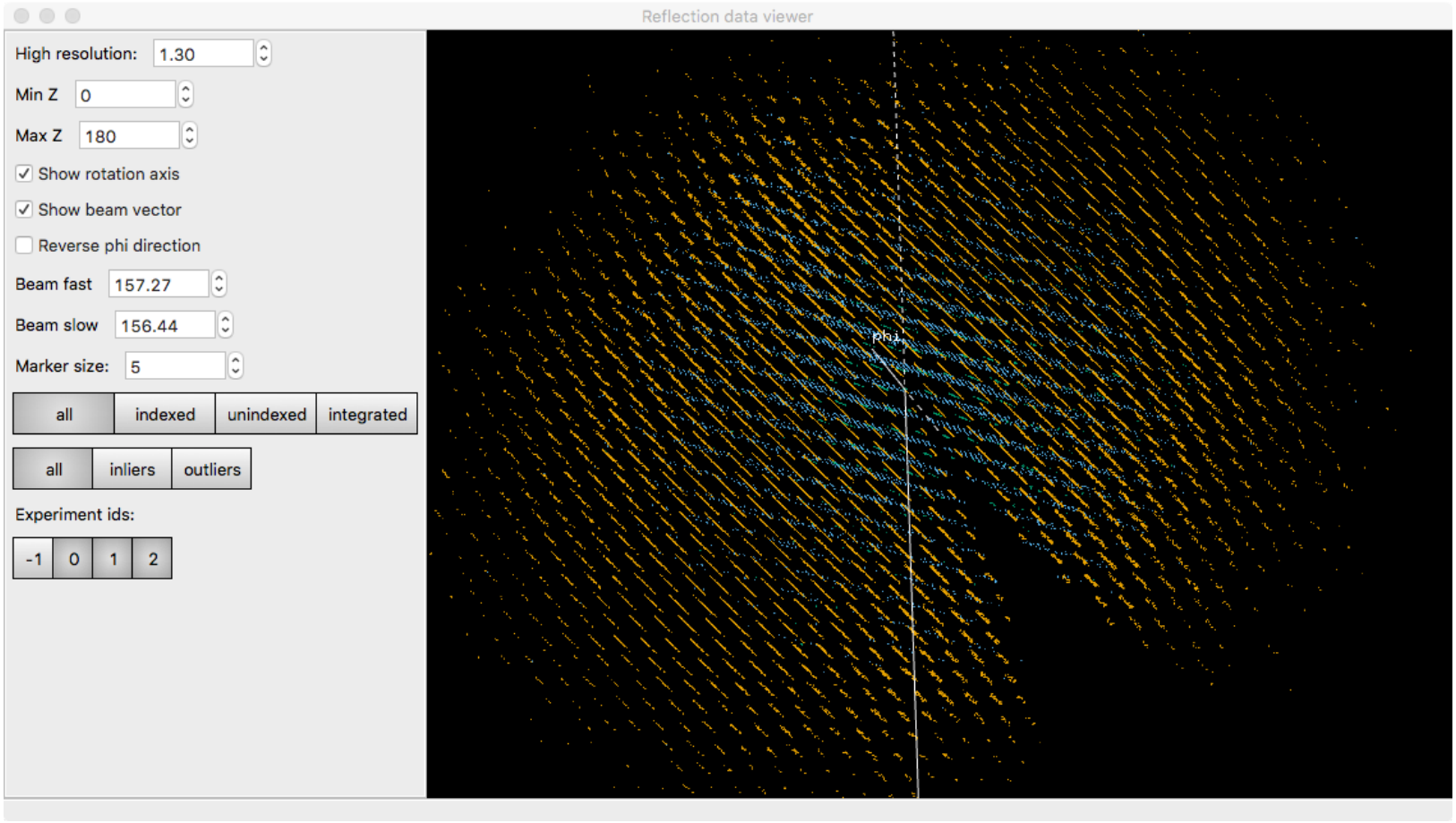




PyMol can be used as well

Sweep of images: spots positions in 3D

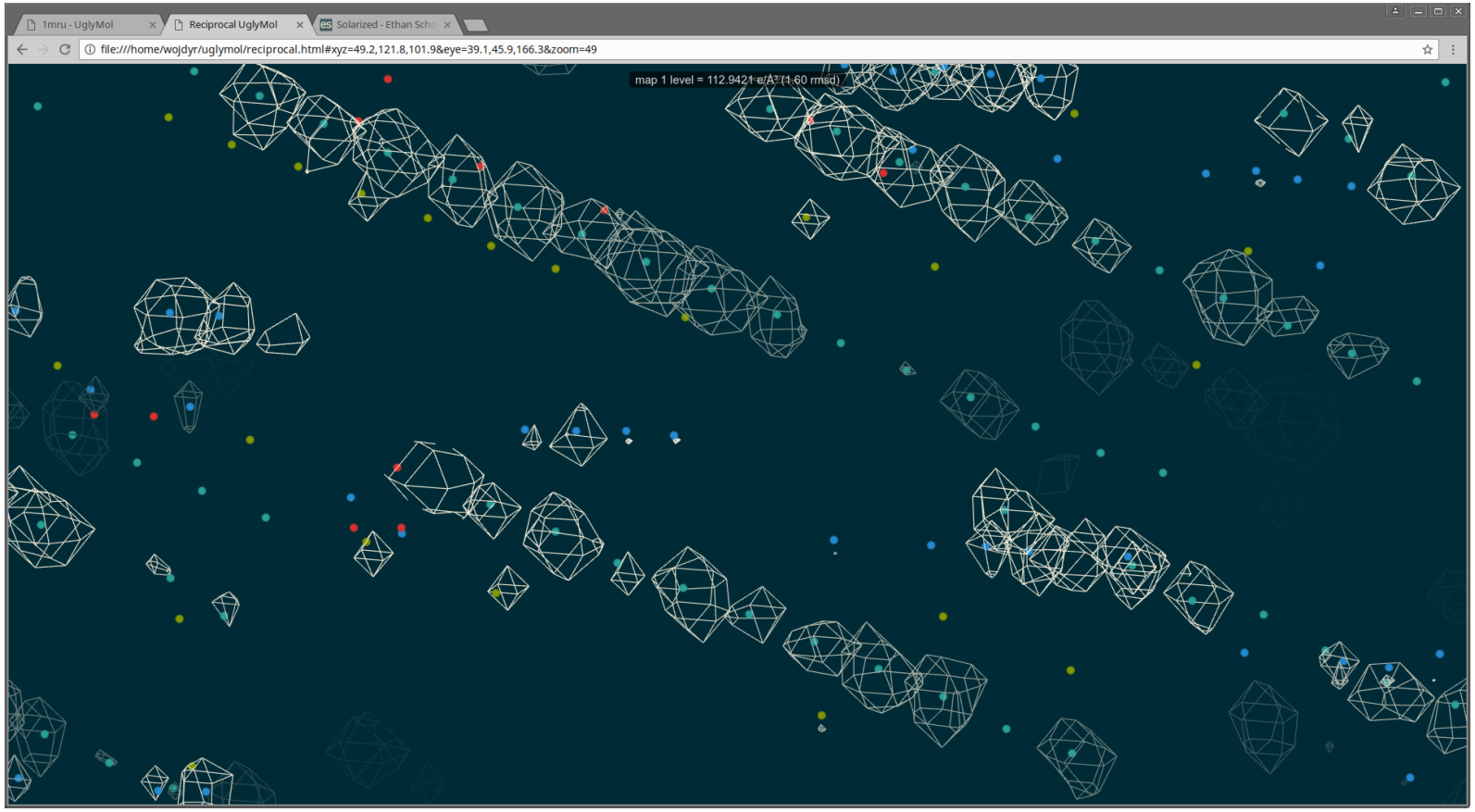
dials.reciprocal_lattice_viewer



Combined views

Combined view, intensities as maps and spots as dots:

- Uglymol in the result page of Xia2 task in CCP4 Cloud
- XDS + spot2pdb + dials.rs_mapper + Coot

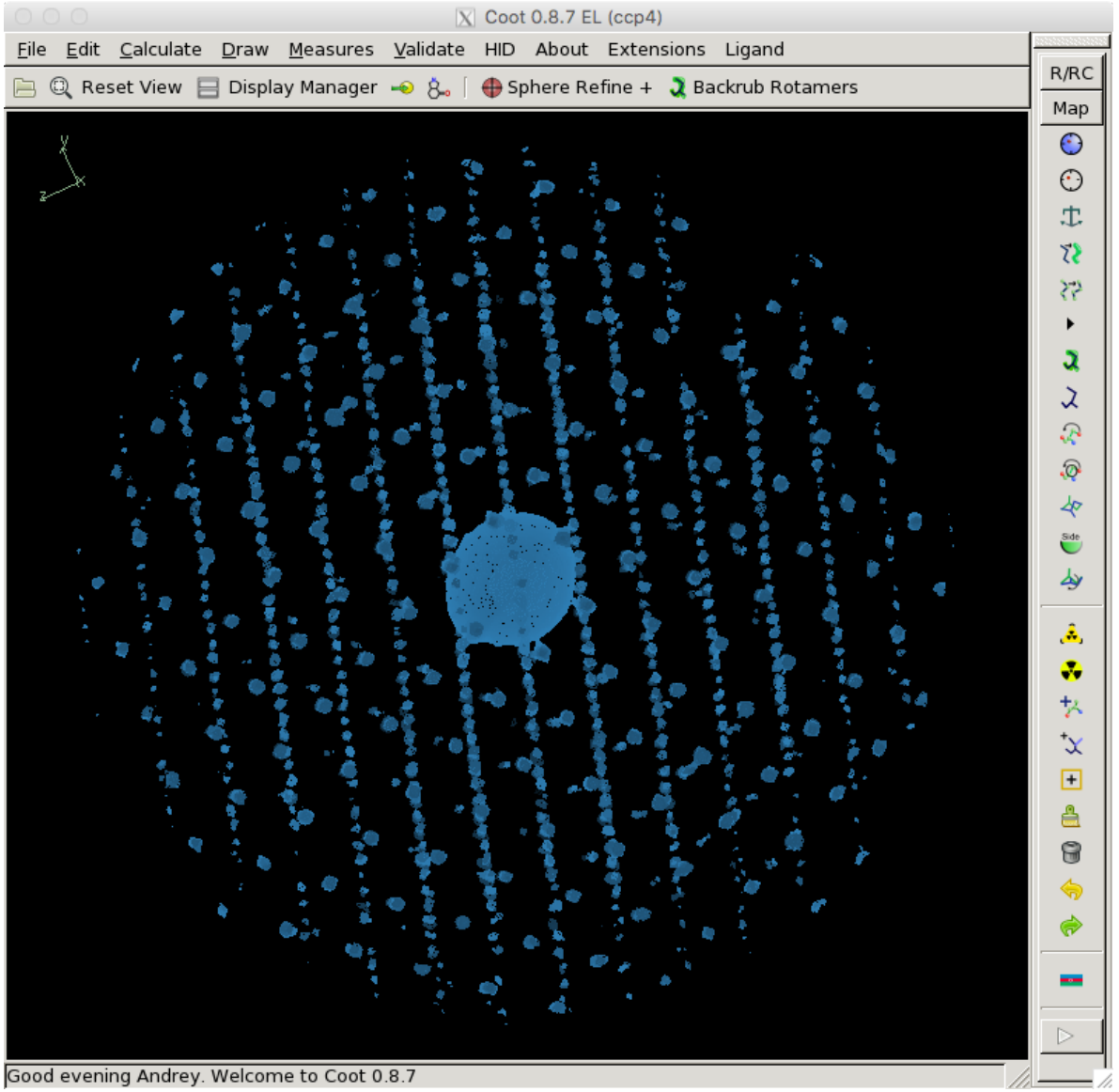


Inter-grown crystals
(multi-lattice data)

Example of random crystal inter-growth

beta-lactamase OXA-163
PDB ID 4s2m

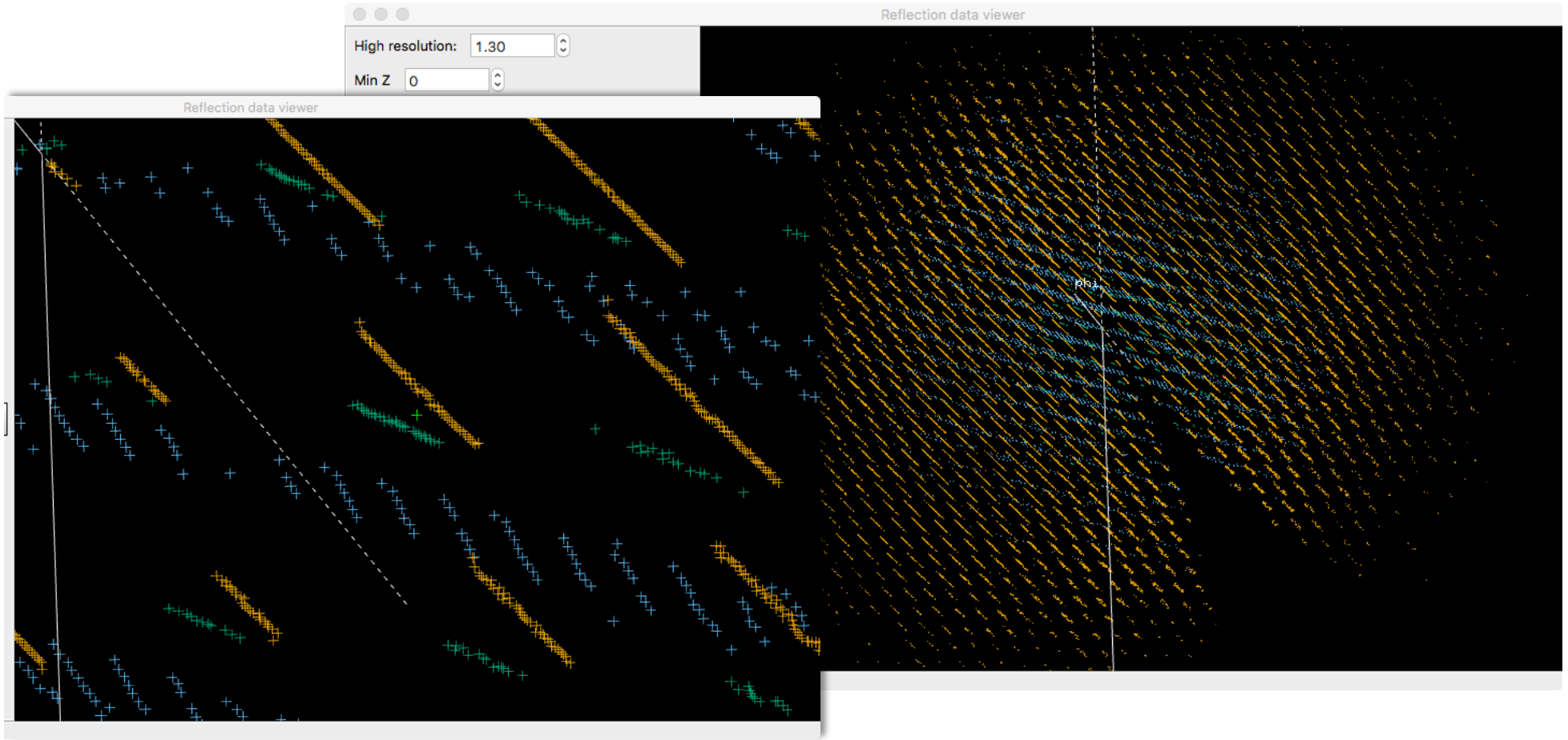
Data from Vlatko Stojanoski
Baylor College of Medicine



Example of random crystal inter-growth

dials.reciprocal_lattice_viewer

- different colour means different lattice
- individual lattices can be switched off and on



Example of random crystal inter-growth

Easy case:

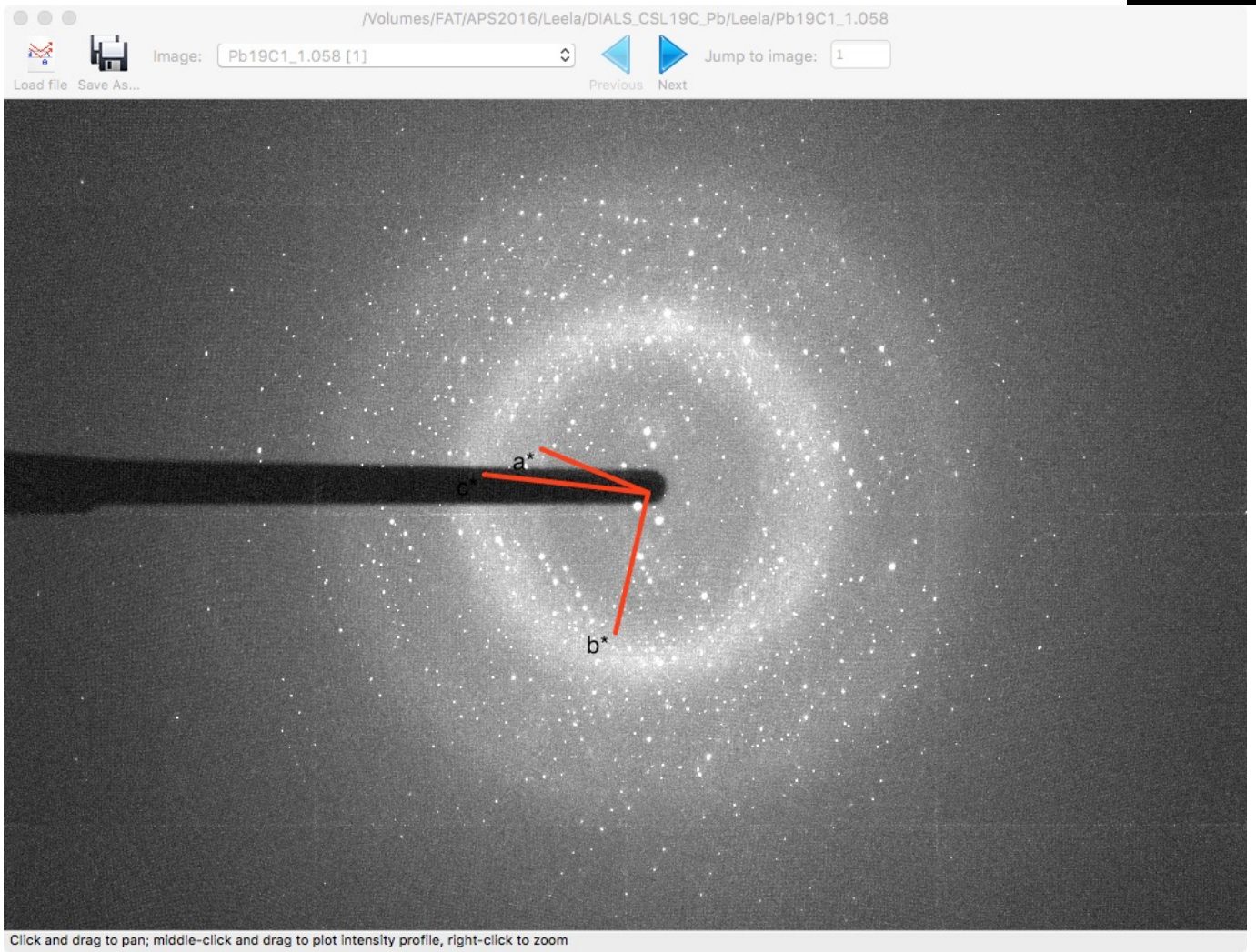
- Lattices are mainly separated, with only very few reflection overlapping
- Signal from one lattice is substantially higher than from others

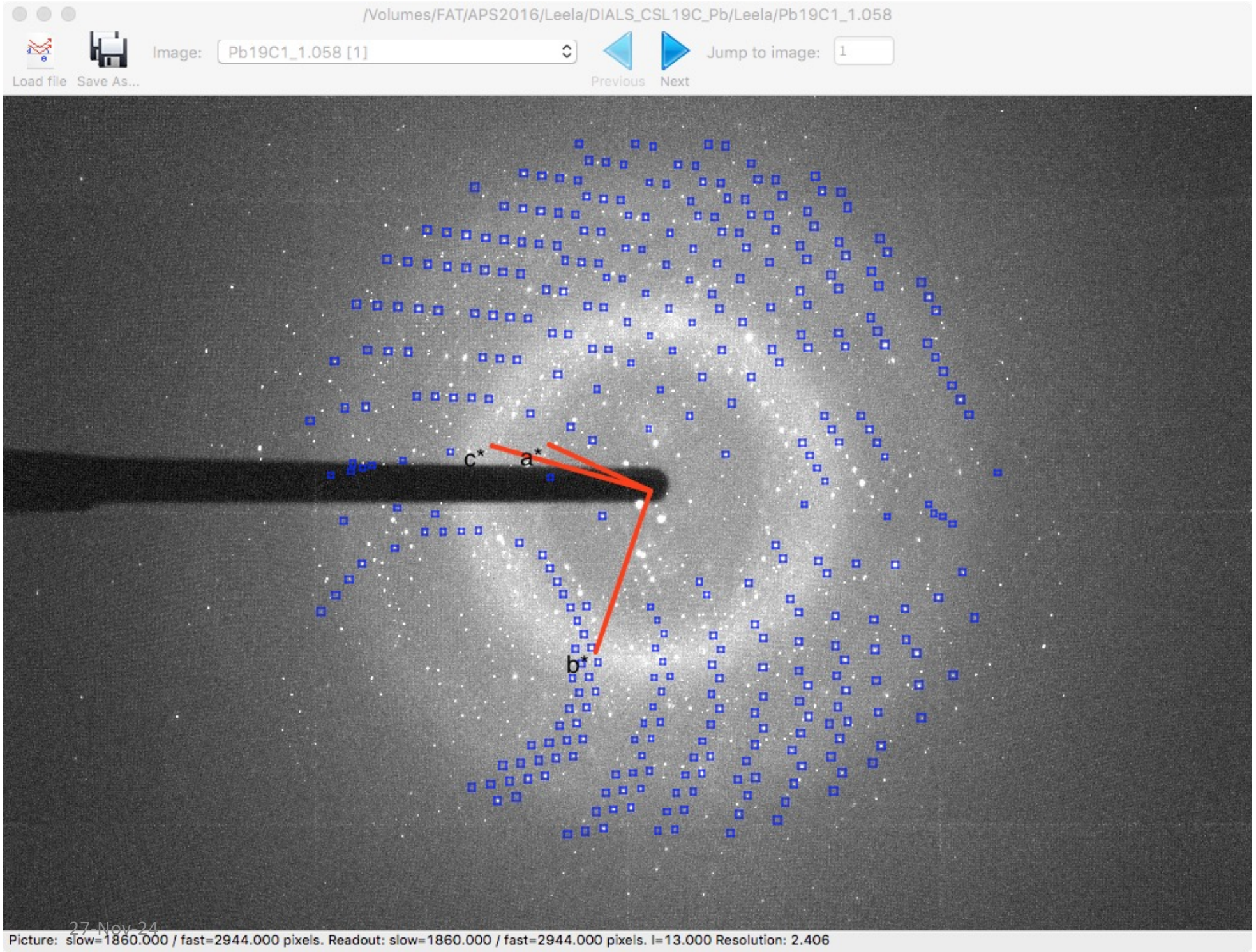
The intensities for the strongest lattice were processed,
structure solved and refined to R=0.20 R-free=0.26

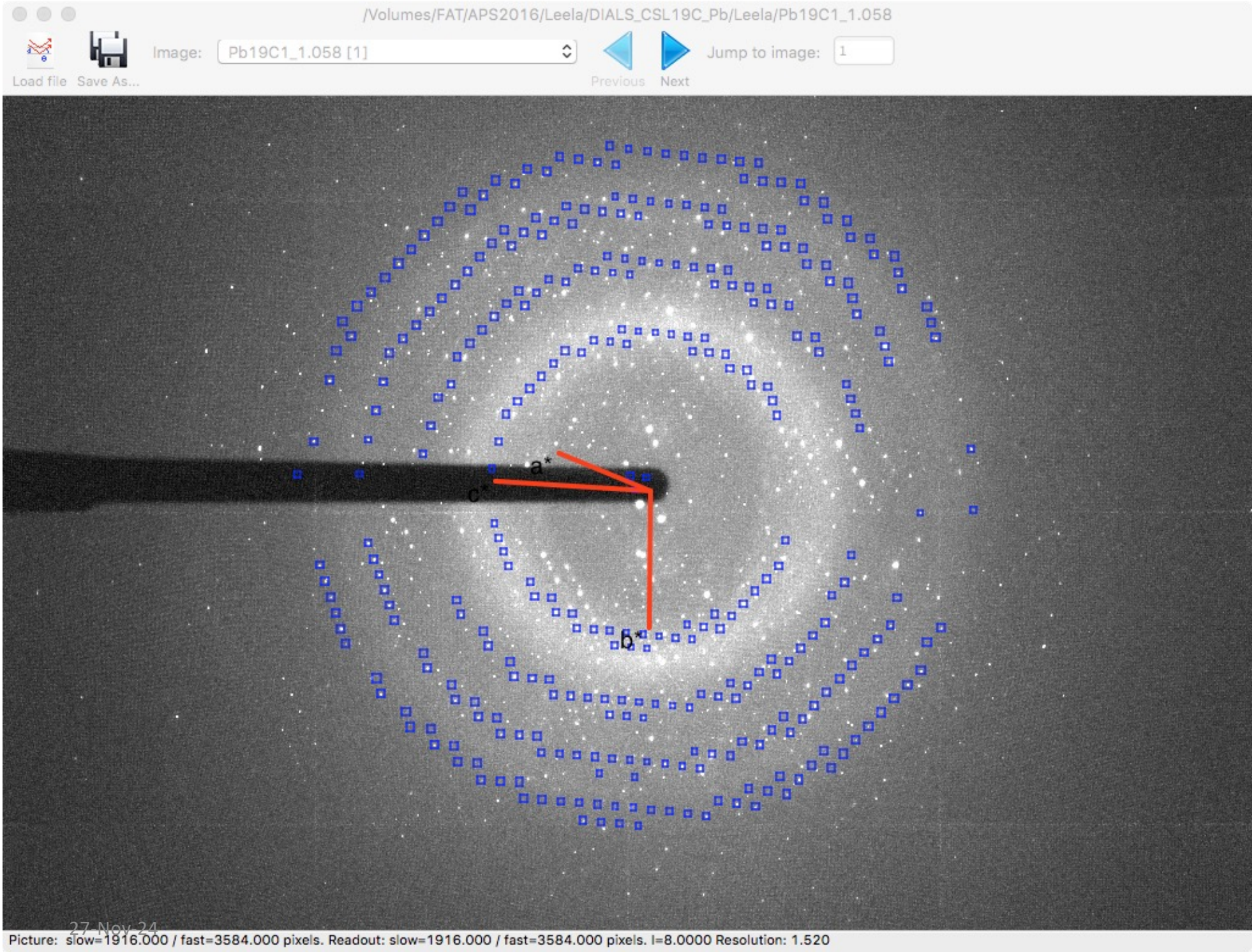
An extreme case

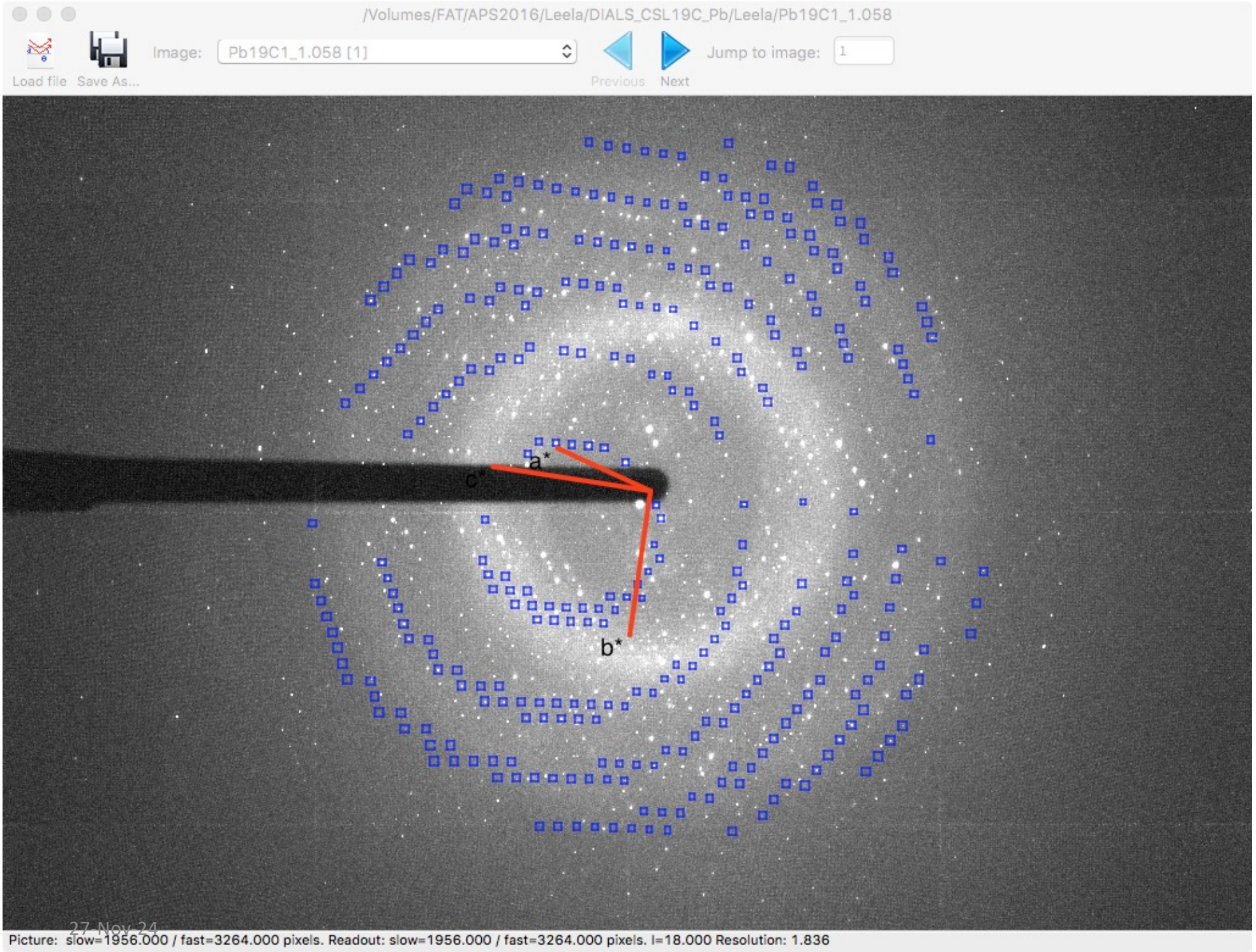
Example from Leela Ruckthong

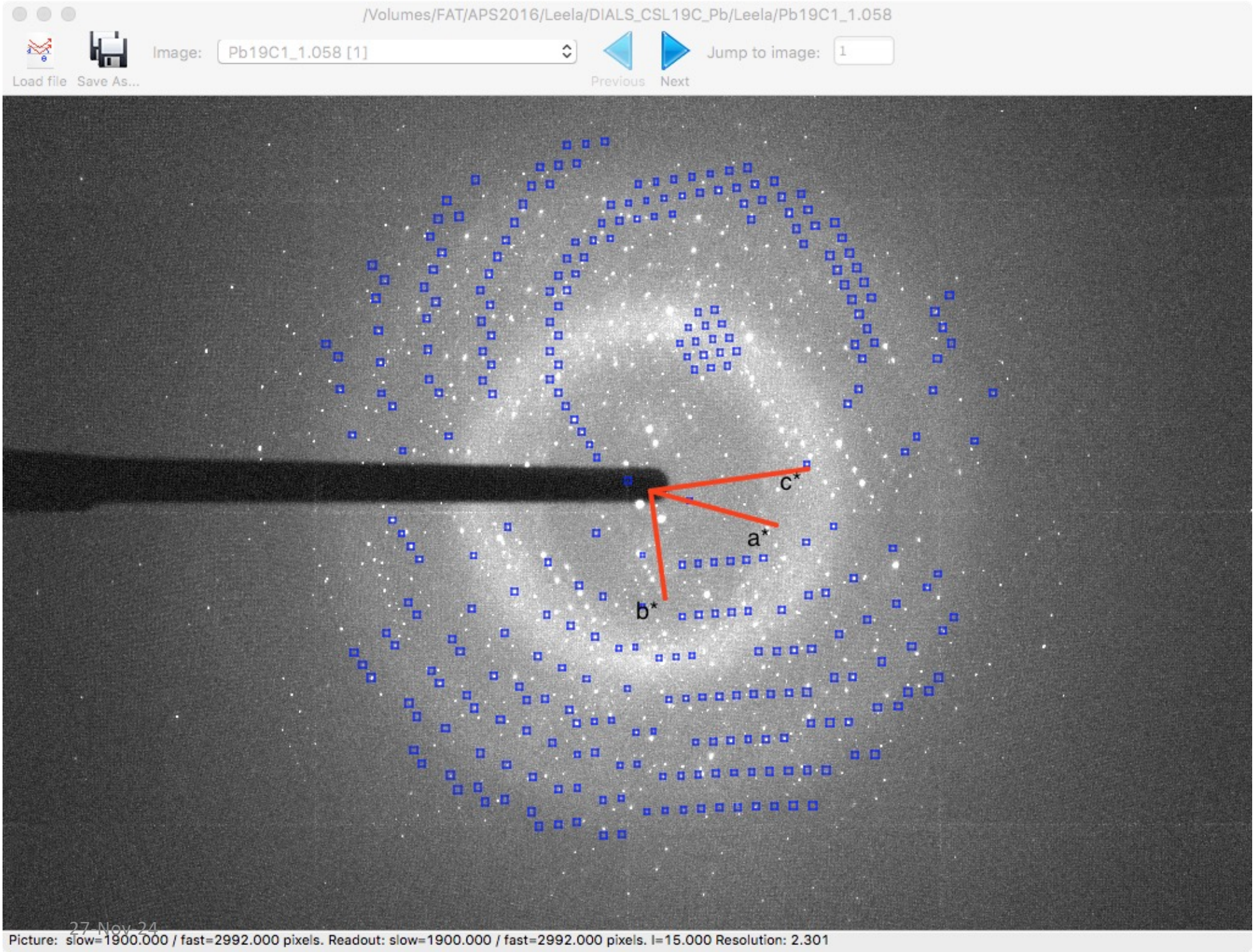
- How many lattices you can spot here?

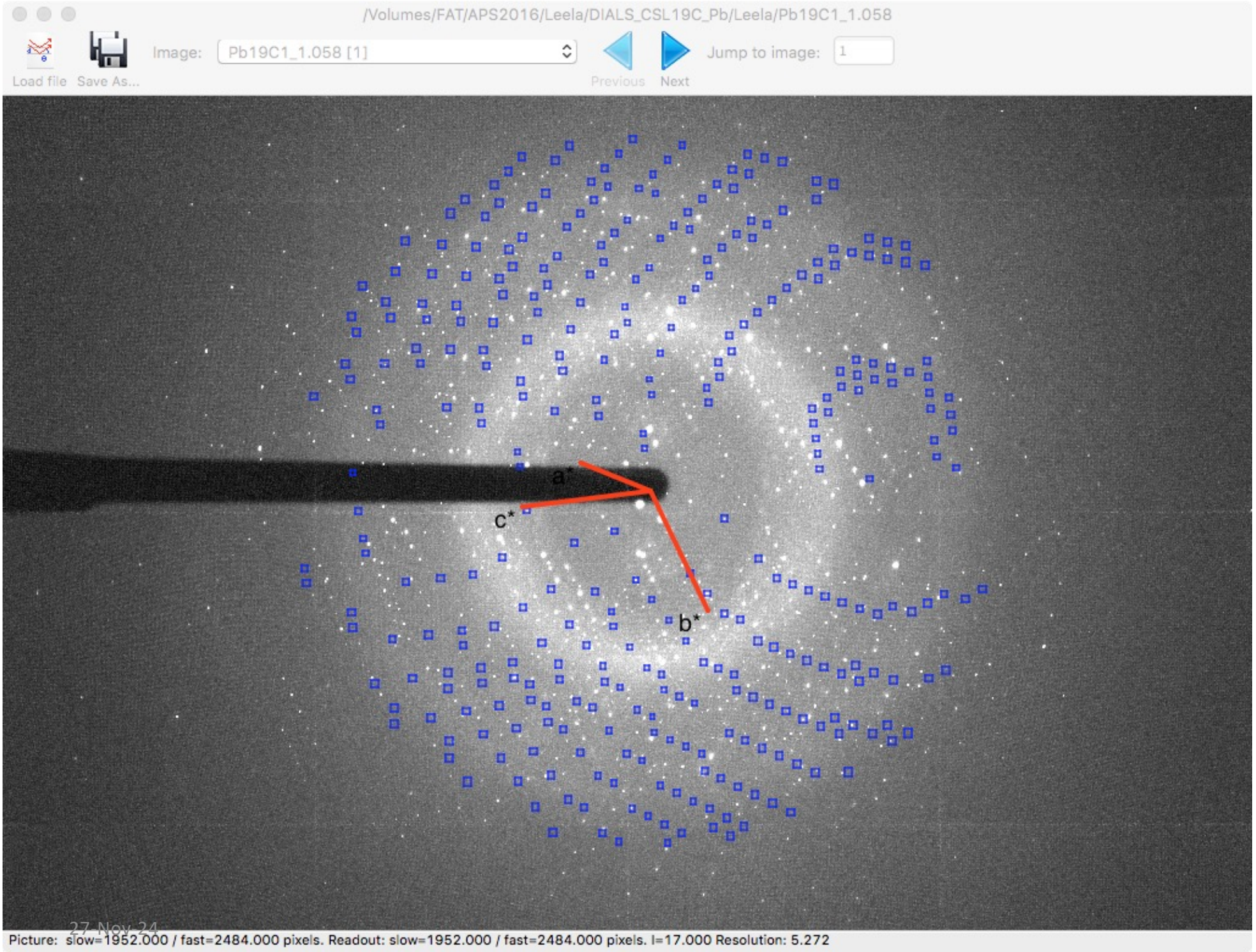


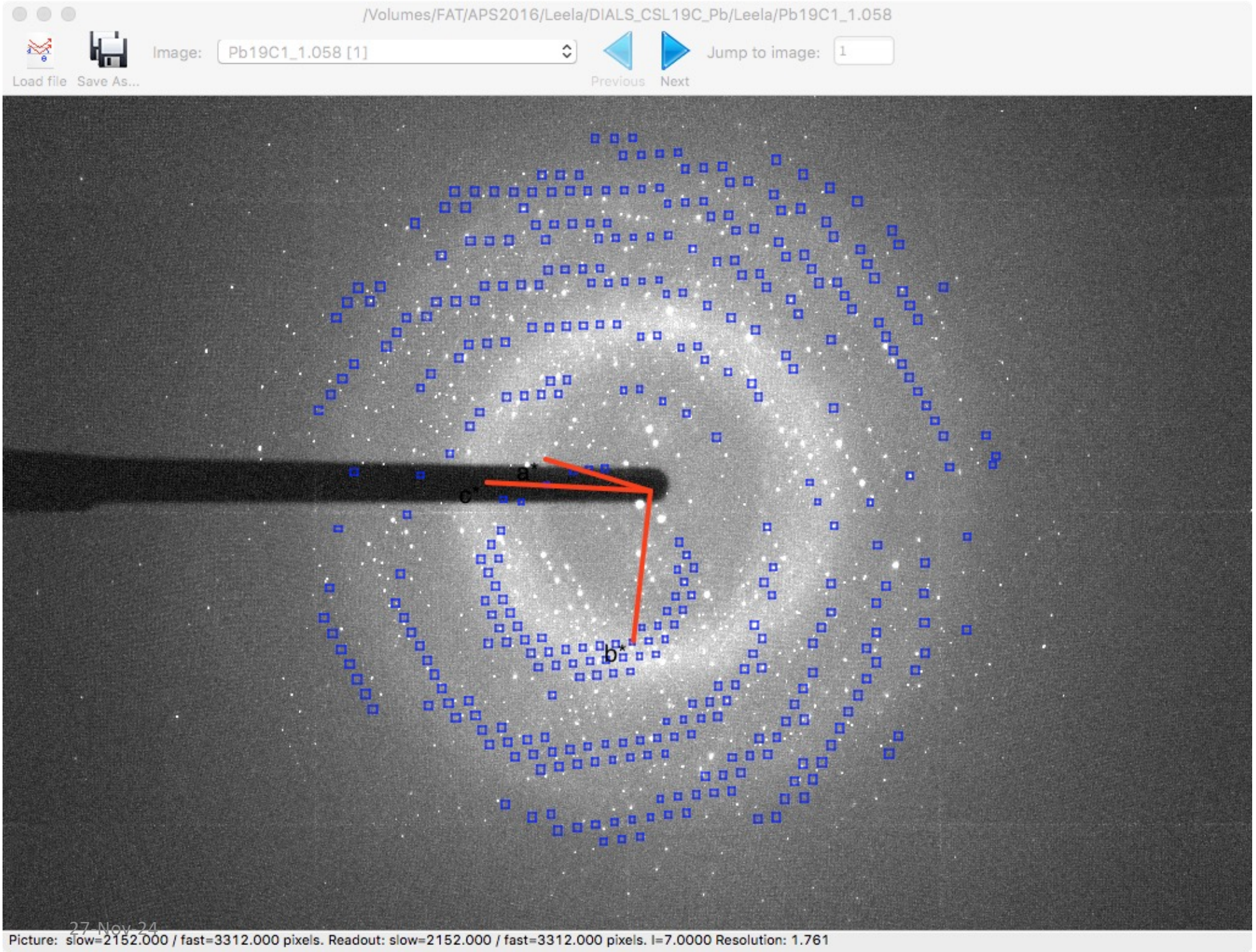

/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.058
Load file
Save As...
Image:
Pb19C1_1.058 [1]
Previous
Next
Jump to image:
1
c*
a*
b*
27-Nov-24
Picture: slow=1860.000 / fast=2944.000 pixels. Readout: slow=1860.000 / fast=2944.000 pixels. I=13.000 Resolution: 2.406

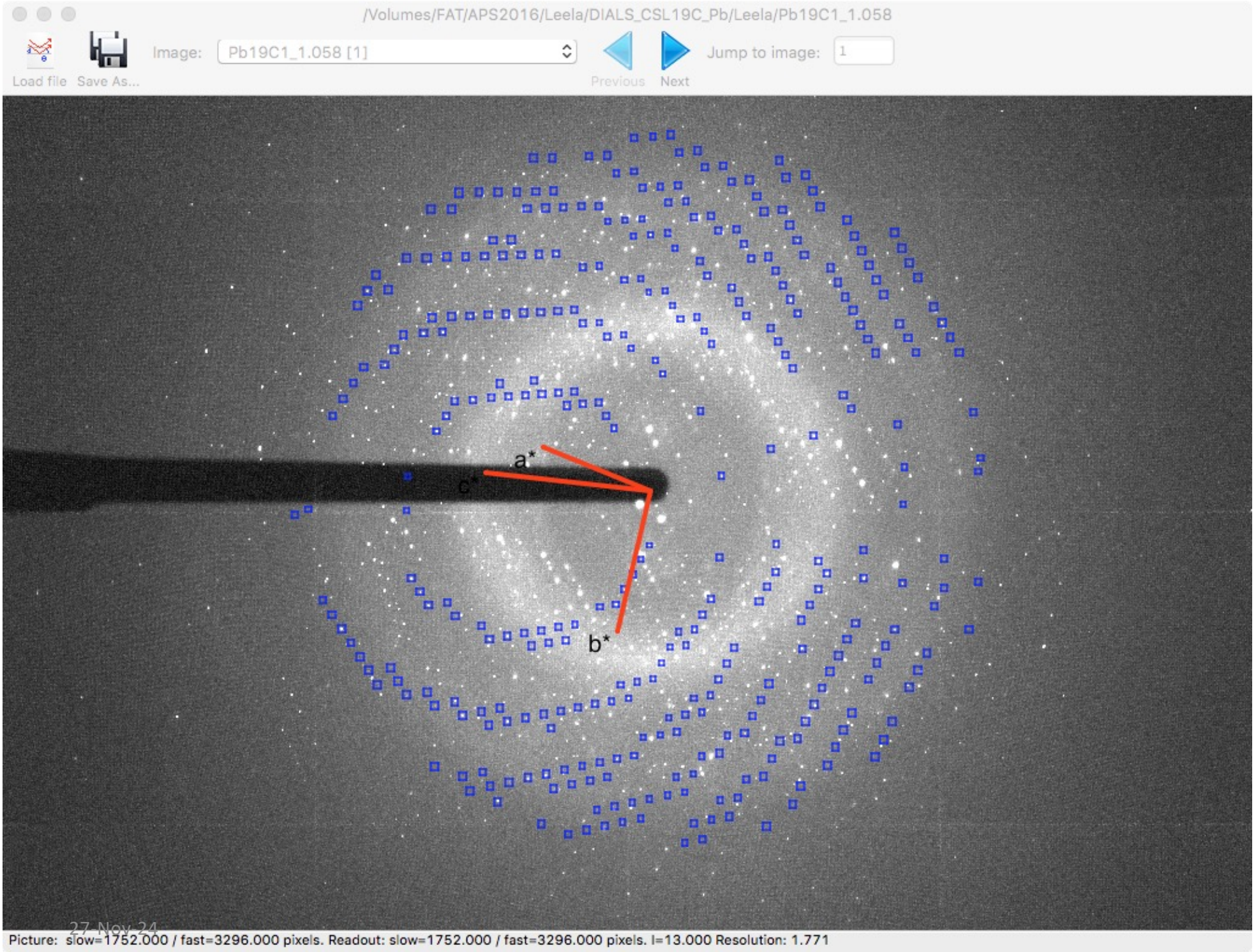

/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.058
Load file
Save As...
Image:
Pb19C1_1.058 [1]
Previous
Next
Jump to image:
1
a*
c*
b*
27-Nov-24
Picture: slow=1916.000 / fast=3584.000 pixels. Readout: slow=1916.000 / fast=3584.000 pixels. I=8.0000 Resolution: 1.520

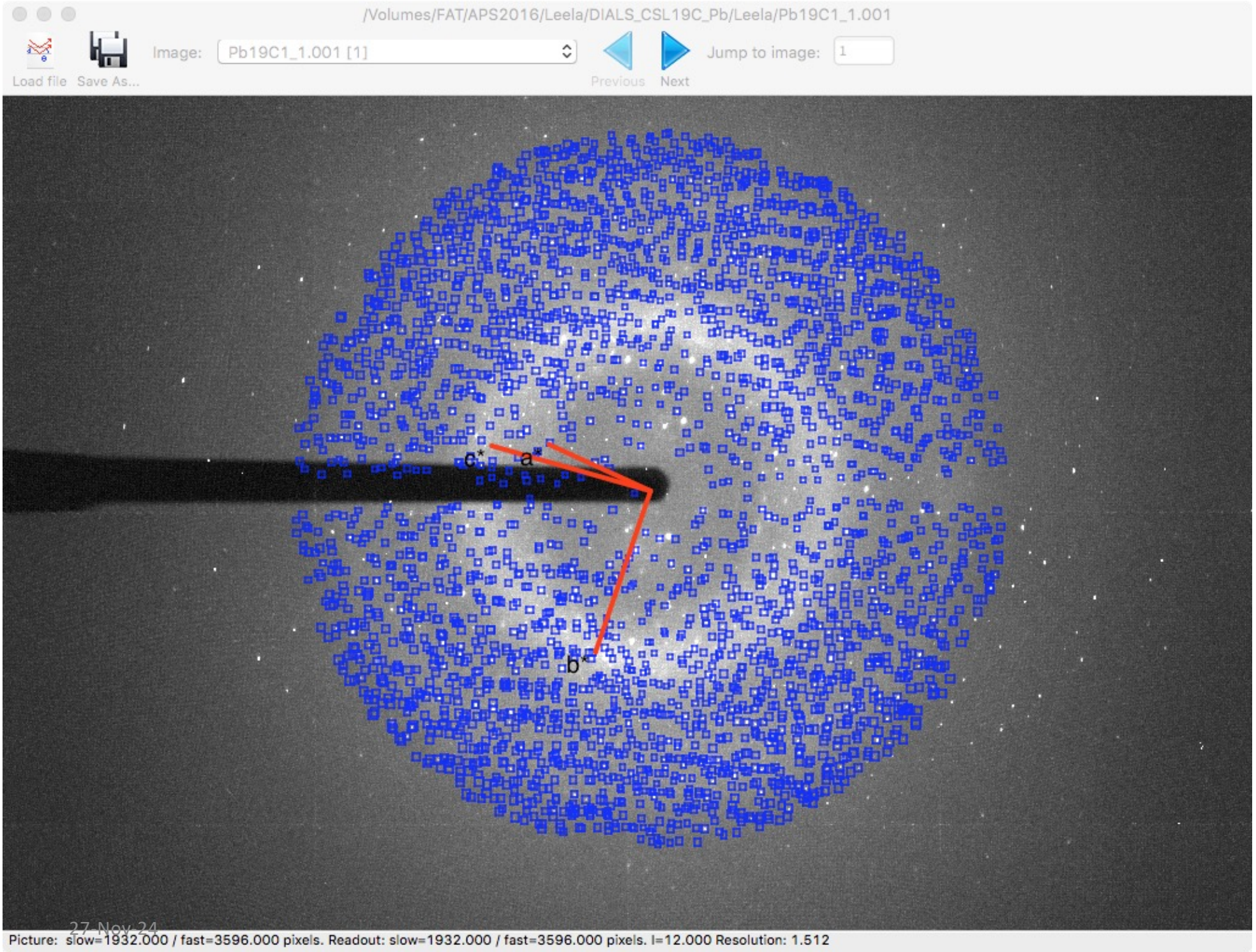

/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.058
Load file
Save As...
Image:
Pb19C1_1.058 [1]
Previous
Next
Jump to image:
1
a*
c*
b*
27-Nov-24
Picture: slow=1956.000 / fast=3264.000 pixels. Readout: slow=1956.000 / fast=3264.000 pixels. I=18.000 Resolution: 1.836

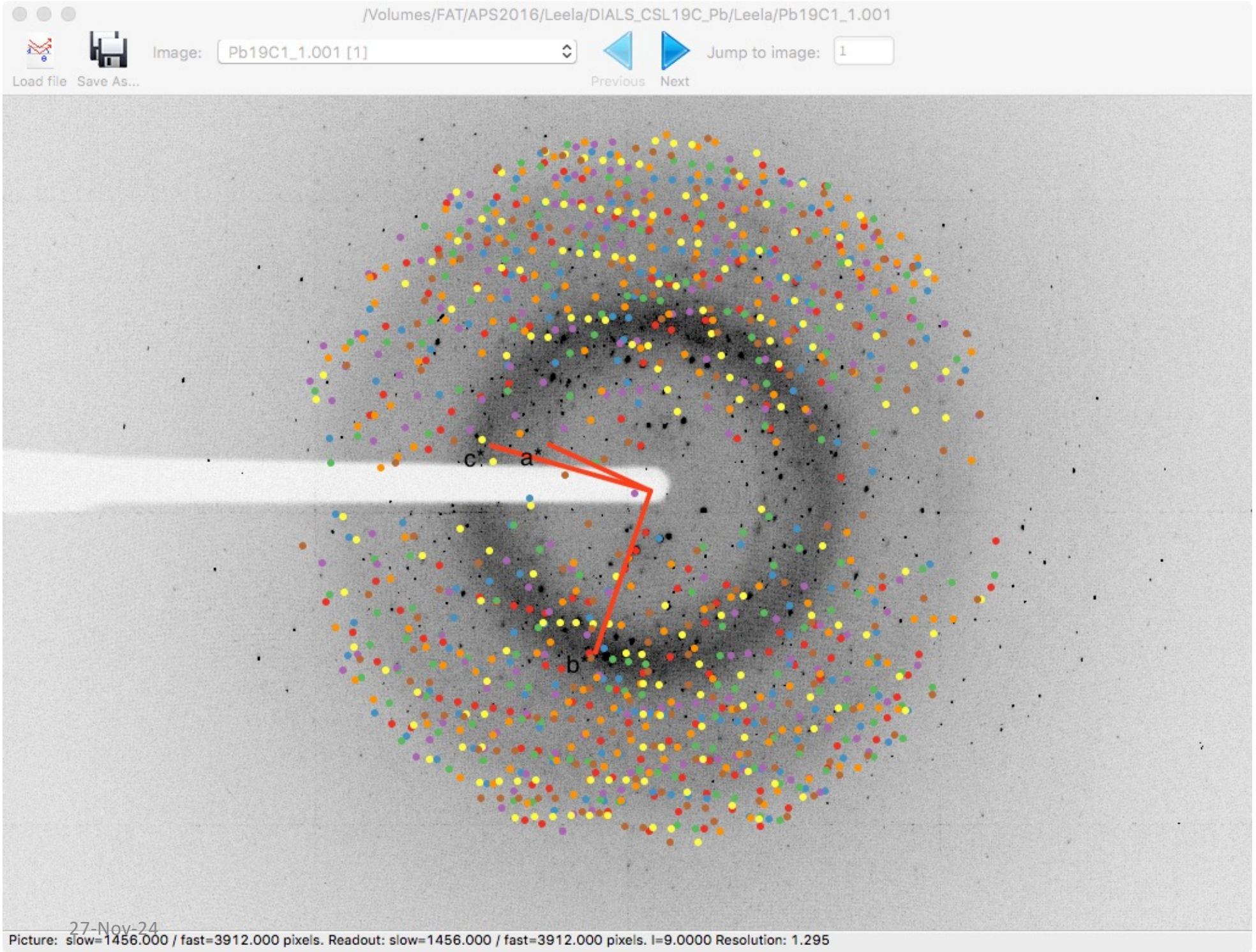

/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.058
Load file
Save As...
Image:
Pb19C1_1.058 [1]
Previous
Next
Jump to image:
1
c*
a*
b*
27-Nov-24
Picture: slow=1900.000 / fast=2992.000 pixels. Readout: slow=1900.000 / fast=2992.000 pixels. I=15.000 Resolution: 2.301


/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.058
Load file
Save As...
Image:
Pb19C1_1.058 [1]
Previous
Next
Jump to image:
1
a*
c*
b*
27-Nov-24
Picture: slow=1952.000 / fast=2484.000 pixels. Readout: slow=1952.000 / fast=2484.000 pixels. I=17.000 Resolution: 5.272


/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.058
Load file
Save As...
Image:
Pb19C1_1.058 [1]
Previous
Next
Jump to image:
1
a*
c*
b*
27-Nov-24
Picture: slow=2152.000 / fast=3312.000 pixels. Readout: slow=2152.000 / fast=3312.000 pixels. I=7.0000 Resolution: 1.761


/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.058
Load file
Save As...
Image:
Pb19C1_1.058 [1]
Previous
Next
Jump to image:
1
a*
c*
b*
27-Nov-24
Picture: slow=1752.000 / fast=3296.000 pixels. Readout: slow=1752.000 / fast=3296.000 pixels. I=13.000 Resolution: 1.771


/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.001
Load file
Save As...
Image:
Pb19C1_1.001 [1]
Previous
Next
Jump to image:
1
c*
a*
b*
27-Nov-24
Picture: slow=1932.000 / fast=3596.000 pixels. Readout: slow=1932.000 / fast=3596.000 pixels. I=12.000 Resolution: 1.512


/Volumes/FAT/APS2016/Leela/DIALS_CSL19C_Pb/Leela/Pb19C1_1.001
Load file
Save As...
Image:
Pb19C1_1.001 [1]
Previous
Next
Jump to image:
1
c*
a*
b*
27-Nov-24
Picture: slow=1456.000 / fast=3912.000 pixels. Readout: slow=1456.000 / fast=3912.000 pixels. I=9.0000 Resolution: 1.295

Seven lattices: too many overlapping spots?

- Only the strongest single lattice gave reasonable merged data
  - » all others were incomplete or had much lower I/sig(I)
  - » merging data from several lattices did not work well
- Solved, refinement stats are not very good

Summary on multiple lattices

- Usually it is reasonable to use the data derived from one singe lattice
- Sometimes completeness can be improved by merging datasets derived from two or more lattices
- Sometimes the second lattice found by Dials result in a dataset of higher resolution than the first one.
- DIALS: Indexing all the lattices together facilitates refinement of the parameters for each individual lattice (consider also xia2.multiplex)

Order-Disorder structures (OD-structures)

- Definition
- Example of an OD twin
- Example of allotwin

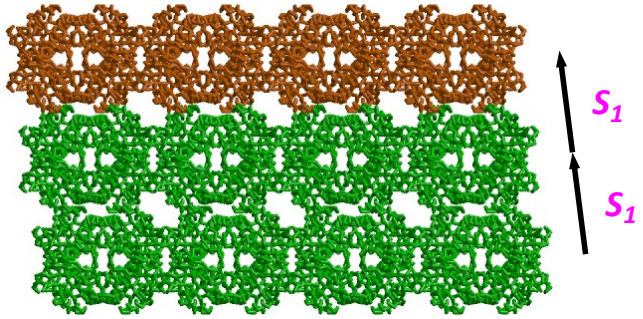
Actually disordered OD-structures:

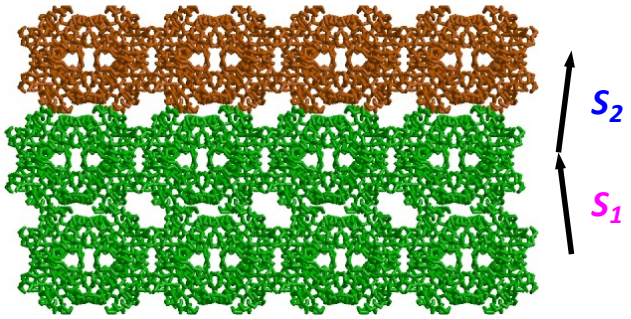
- Relatively frequent
- Good for illustration purpose

Order-disorder structures (OD-structures)

- identical layers
- identical interfaces between the layers
- but: two or more ways of packing three adjacent layers

*) MX: "identical" means Ca r.m.s.d. < 1 A



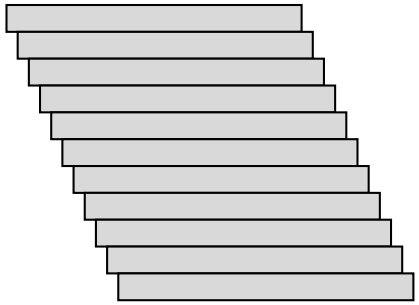


*) S_1 and S_2. are called stacking vectors

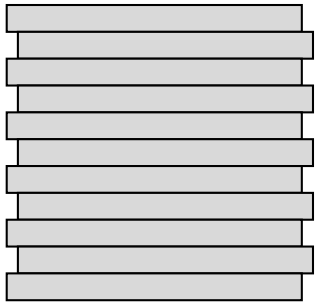
- two-dimensional periodicity
- a potential for disorder in the third dimension

OD-structures

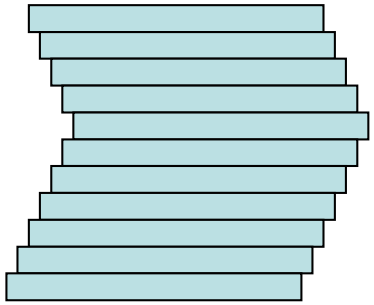
Single crystal



Single crystal

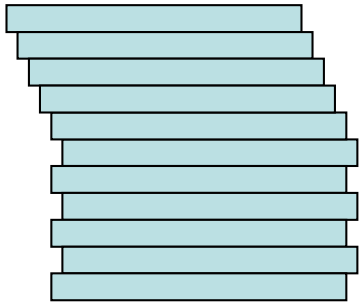


OD-twin



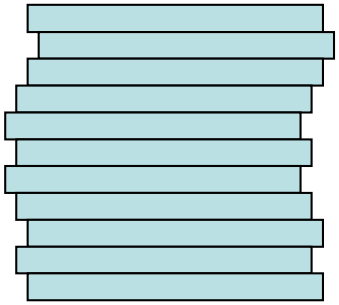
Example 1

Allotwin



Example 2

Partially disordered OD-structure



Examples in the next section

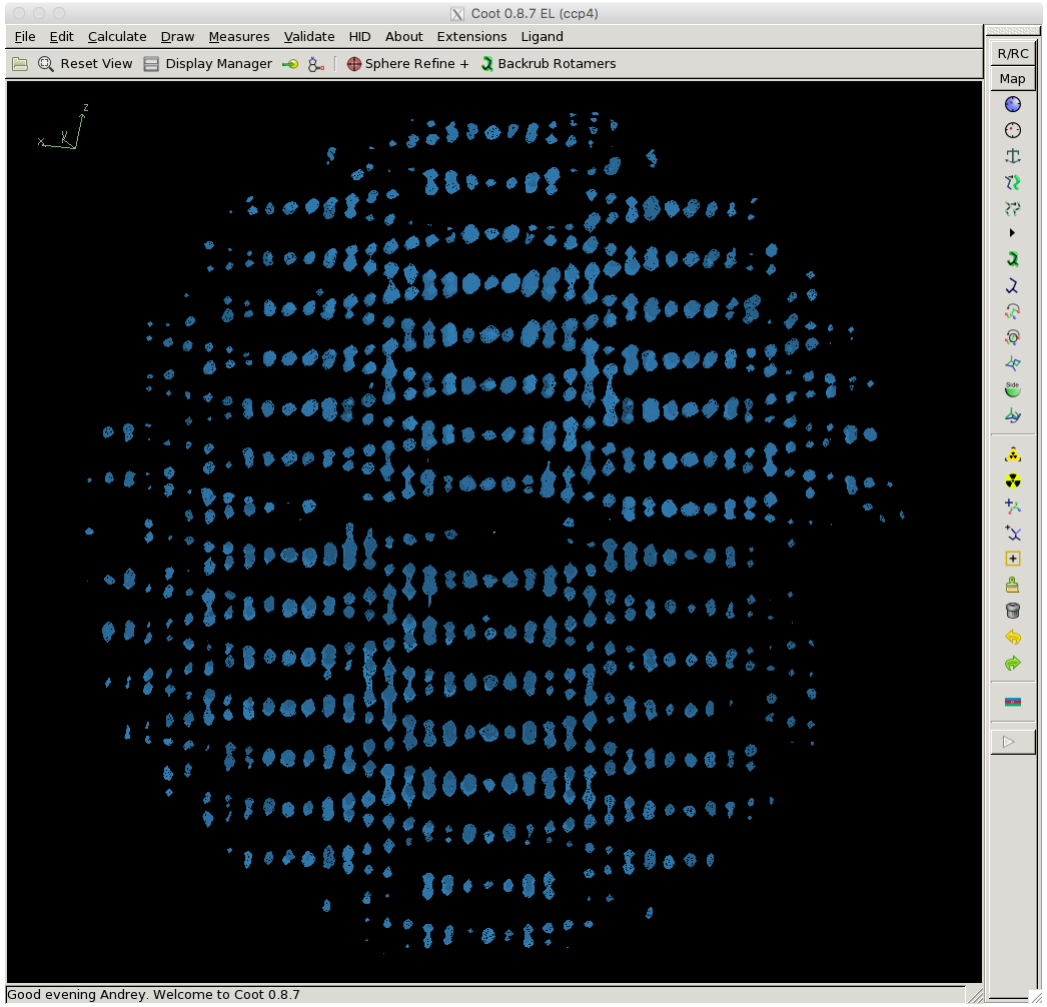
Example 1: OD-twin

Morphological classification
OD-twin

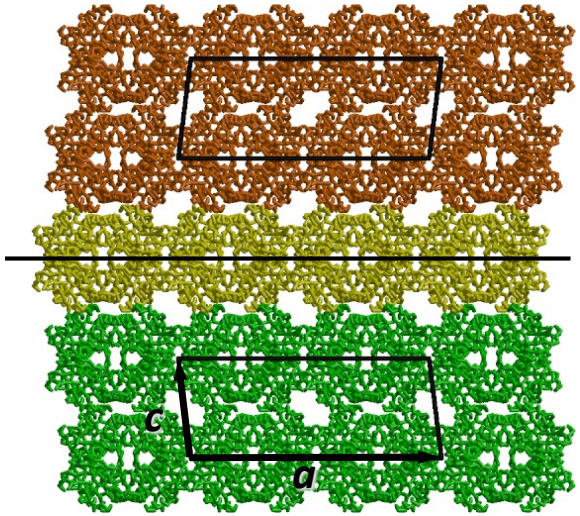
Geometrical classification
Twinning by reticular pseudo-merohedry

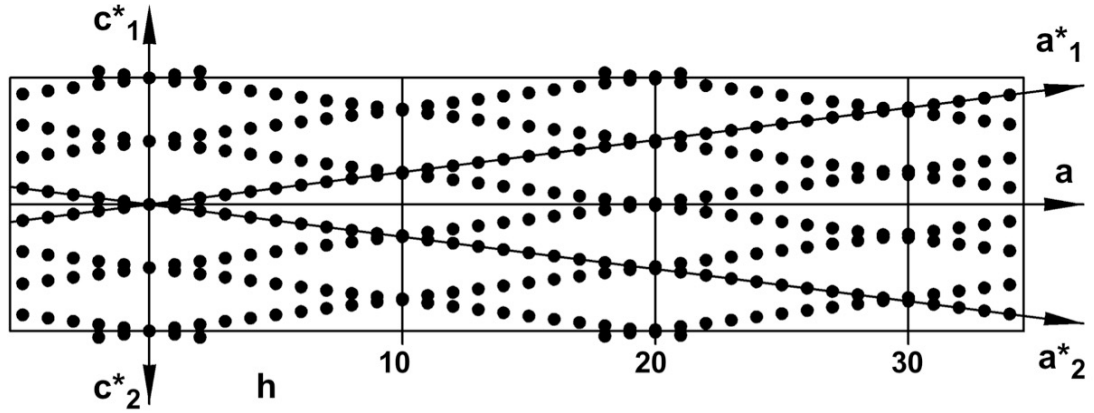
Synonym
Non-merohedral twinning

dials.rs_mapper + coot



Real and reciprocal lattices



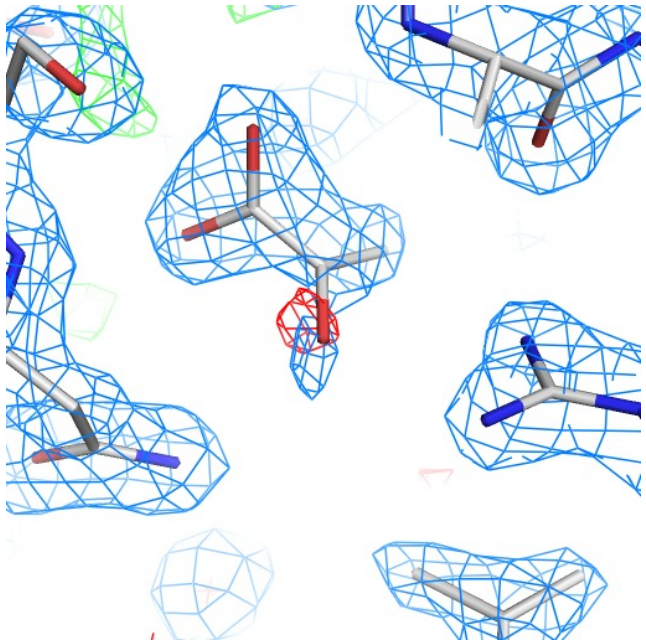


- Process data from one lattice and ignore twinning
- Process data from one lattice and demodulate the data

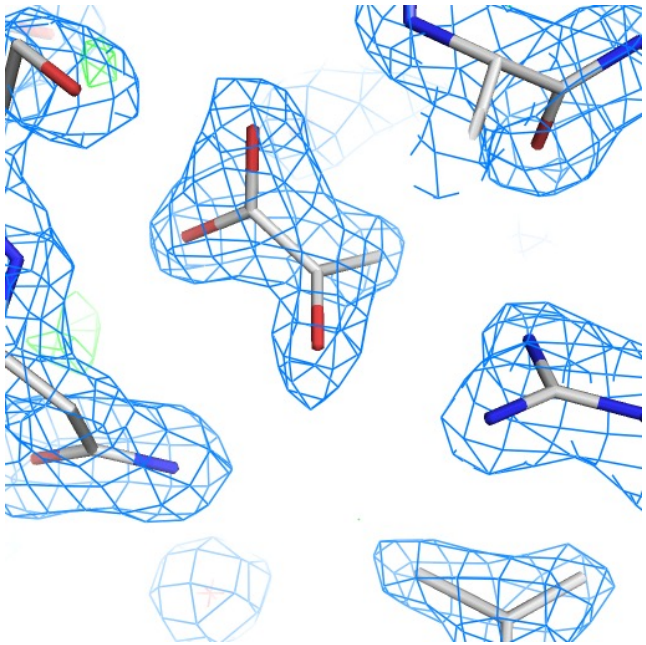
- Record total intensity of overlapping spots (SAINT, iMosflm) and deal with it at refinement (SHELXL)

Demodulation: improvement in the electron density

Visually, improvement from demodulation occurred
only in the electron density for solvent molecules (here: L-lactate)

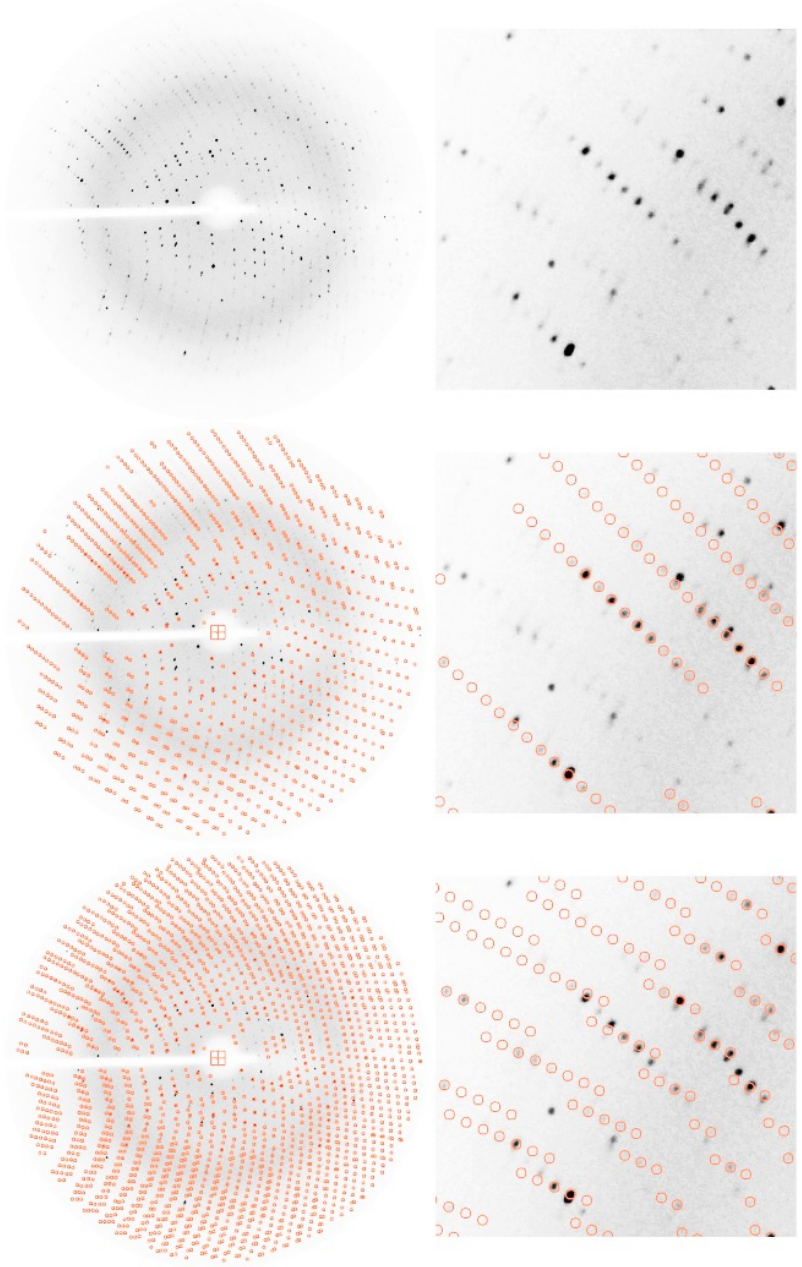


Before demodulation
R / R-free = 0.21 / 0.27

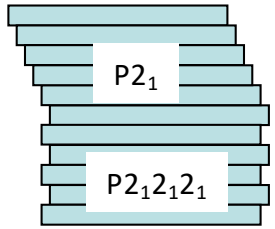


After demodulation
R / R-free = 0.16 / 0.23

Example 2: allotwin



Crystals of Lon protease
Resolution 3Å



Dauter *et al.* (2005).
Acta Cryst. D**61**, 967-975.

$P2_1$

a = 48.5 Å
b = 86.3 Å
c = 138.0 Å
β = 92.3°

$P2_12_12_1$

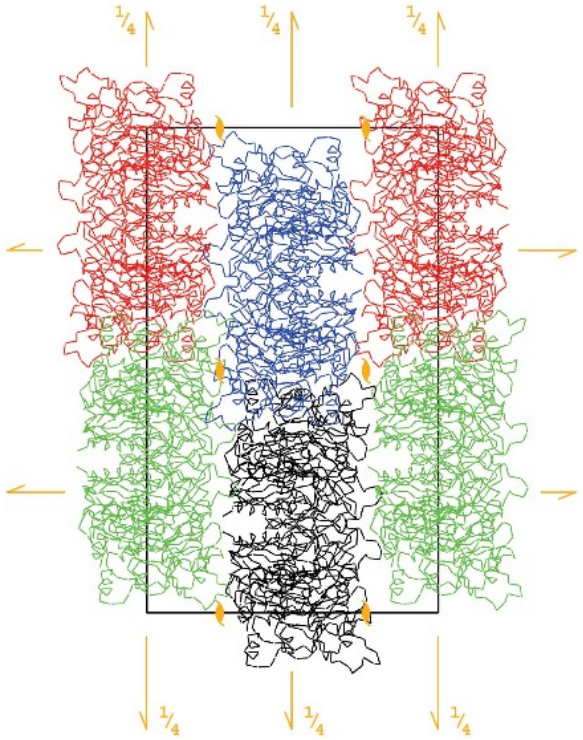
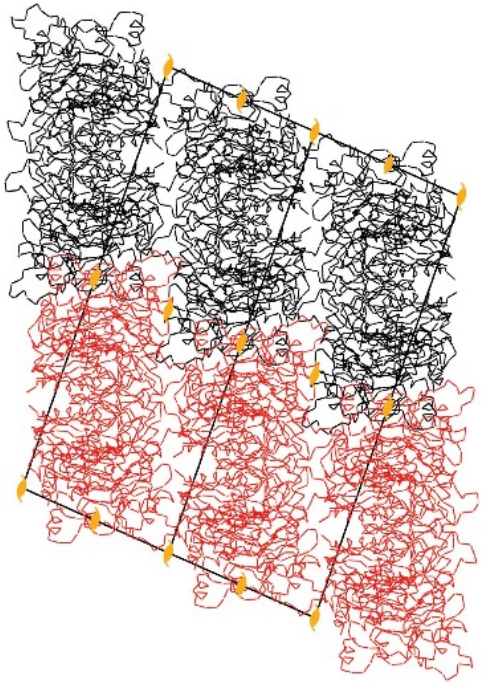
a = 86.3 Å
b = 90.6 Å
c = 148.0 Å

Example 2: allotwin

Crystals of Lon protease
Resolution 3Å

Dauter *et al.* (2005).
Acta Cryst. D**61**, 967-975.

Structures of both crystal forms were solved from a one allotwin crystal

	PDB code 1z0t	PDB code 1z0v
		
	$P2_12_12_1$	$P2_1$
R / R-free	0.19 / 0.35	0.21 / 0.31

Example 2: allotwin

- More frequently, the presence of very different indexing solutions means that the indexing program is struggling rather than there are domains belonging to different space groups.
- 3D viewers will help to check what is actually happening.
- In any case, it is a warning sign if e.g. different programs suggest space groups belonging to different point groups (e.g. *C* 2 and *P* $2_12_12_1$)

Example from ISPyB

Auto Processing — xia2 dials: ✔

Type	Resolution	Spacegroup	Mn<I/sig(i)>	Rmeas Inner	Rmeas Outer	Completeness	
xia2 dials	68.91 - 2.54	C 1 2 1	3.9	0.147	3.639	99.6	137.98 29.8
xia2 3dii	56.29 - 2.78	P 21 21 21	4.7	0.149	2.577	100.0	29.93 112.5
autoPROC							
fast_dp							
2x xia2.multiplex	92.39 - 2.23	C 1 2 1	4.4	0.146	5.270	100.0	137.70 29.8

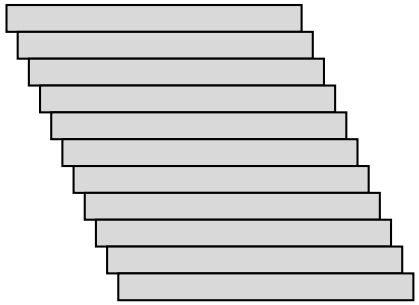
xia2 dials | xia2 3dii | autoPROC | fast_dp | 2x xia2.multiplex

Partially disordered OD-structures

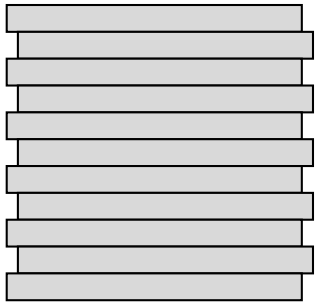
- Visualisation (detection)
- Ghost density
- Indexing
- Effect on structure solution

OD-structures

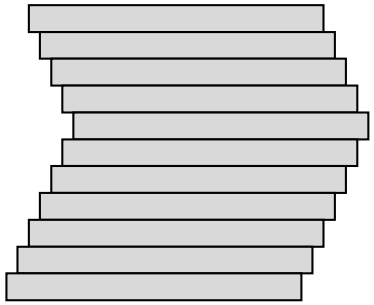
Single crystal



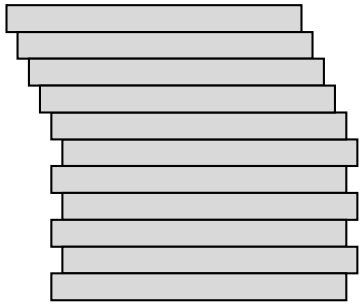
Single crystal



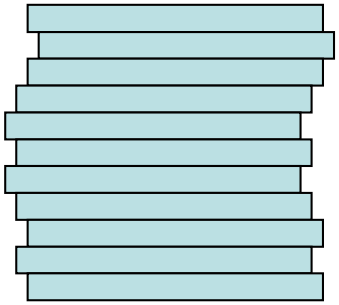
OD-twin



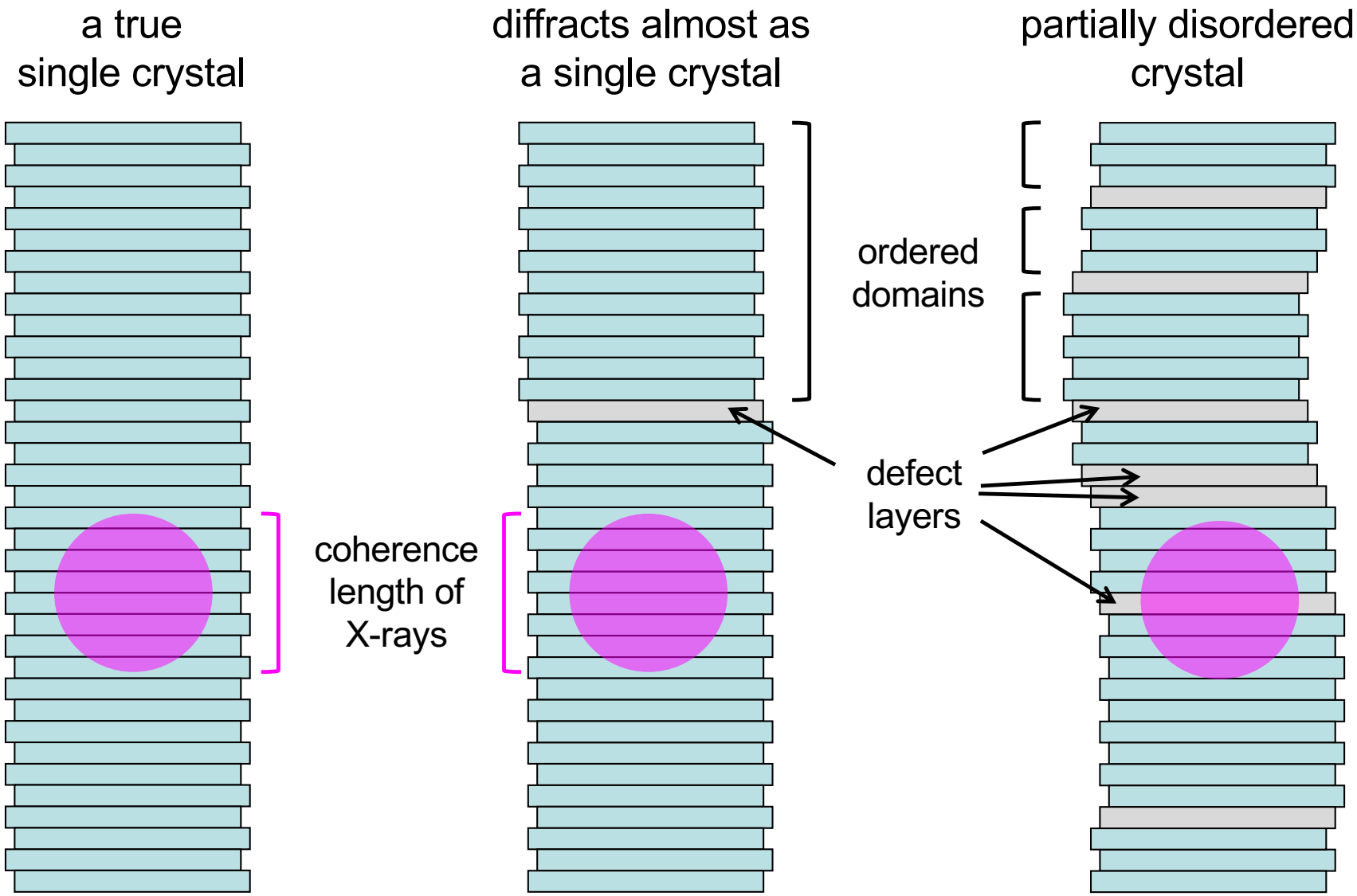
Allotwin



Partially disordered OD-structure

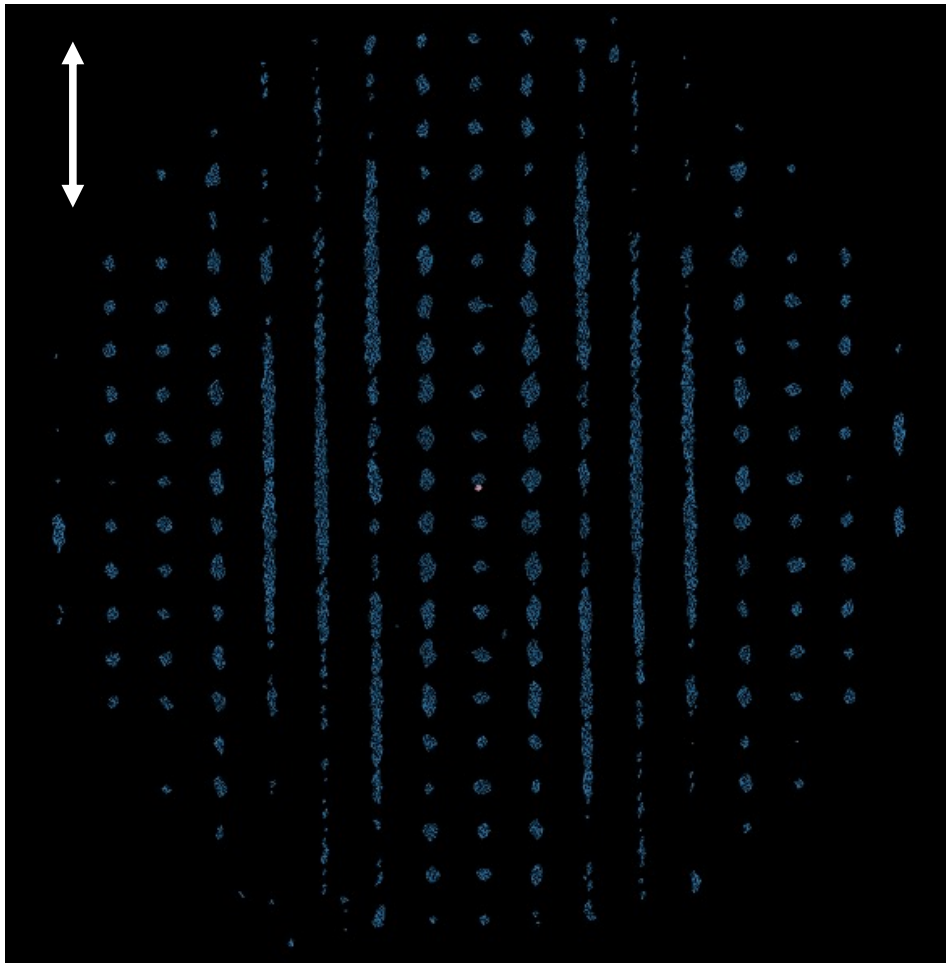


Partially disordered OD structures

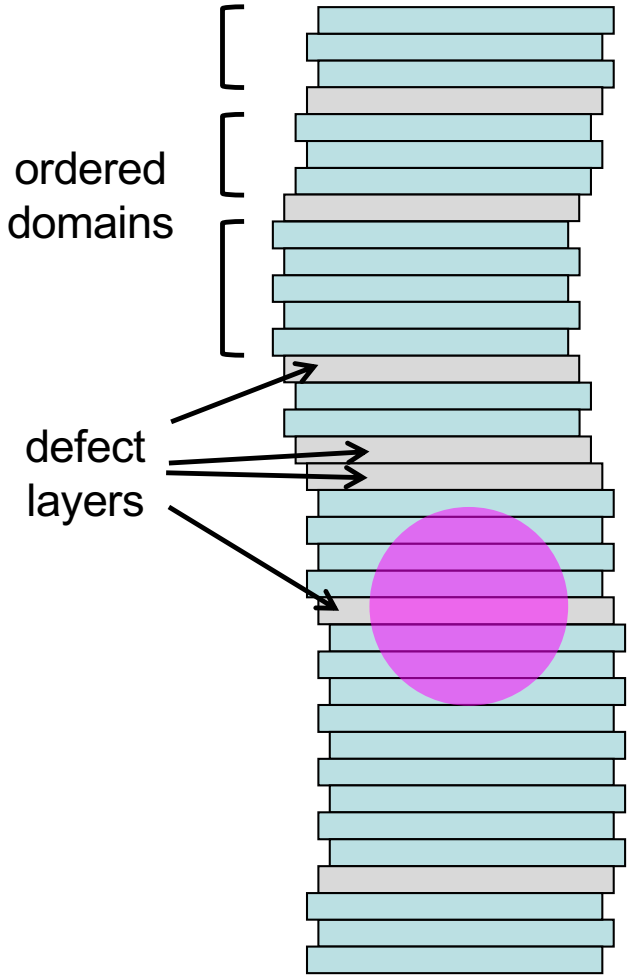


Diffraction of partially disordered structures

White arrow - direction in which global periodicity is missing

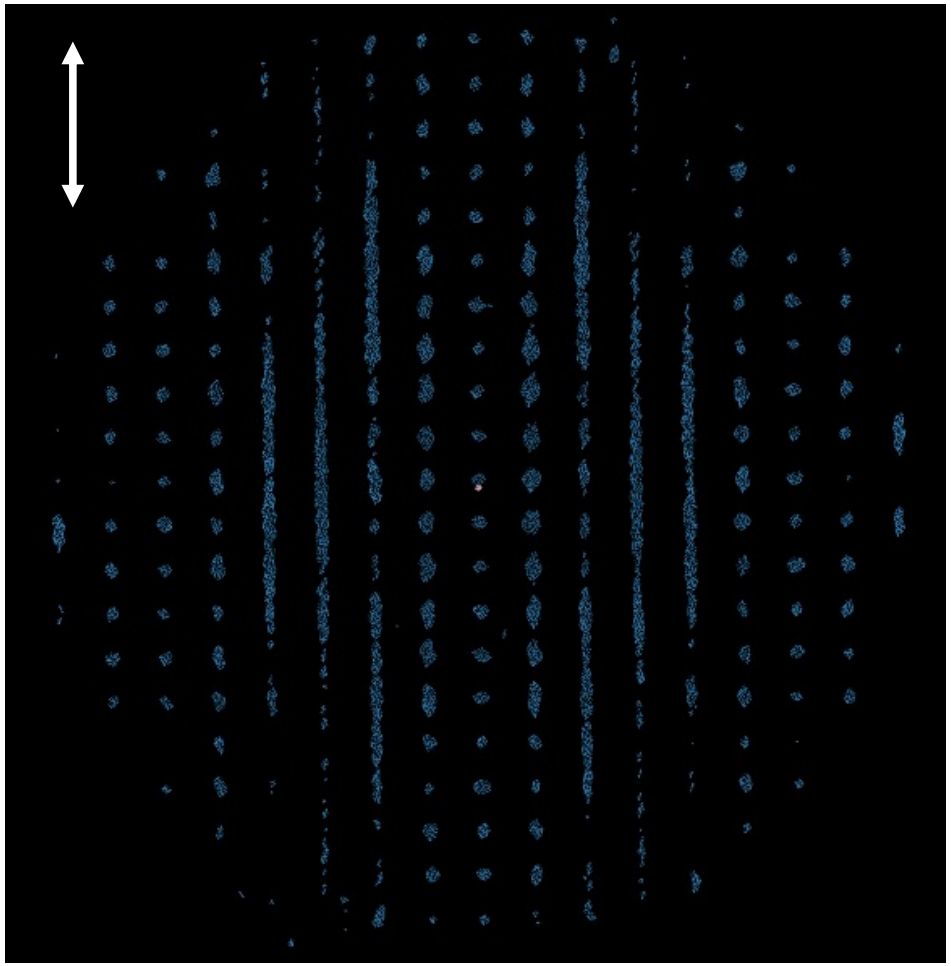


dials.rs_mapper + coot



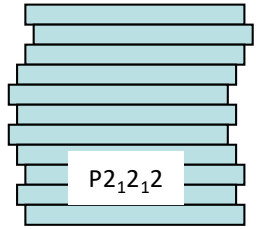
Example 1: ghost density

White arrow - direction in which global periodicity is missing



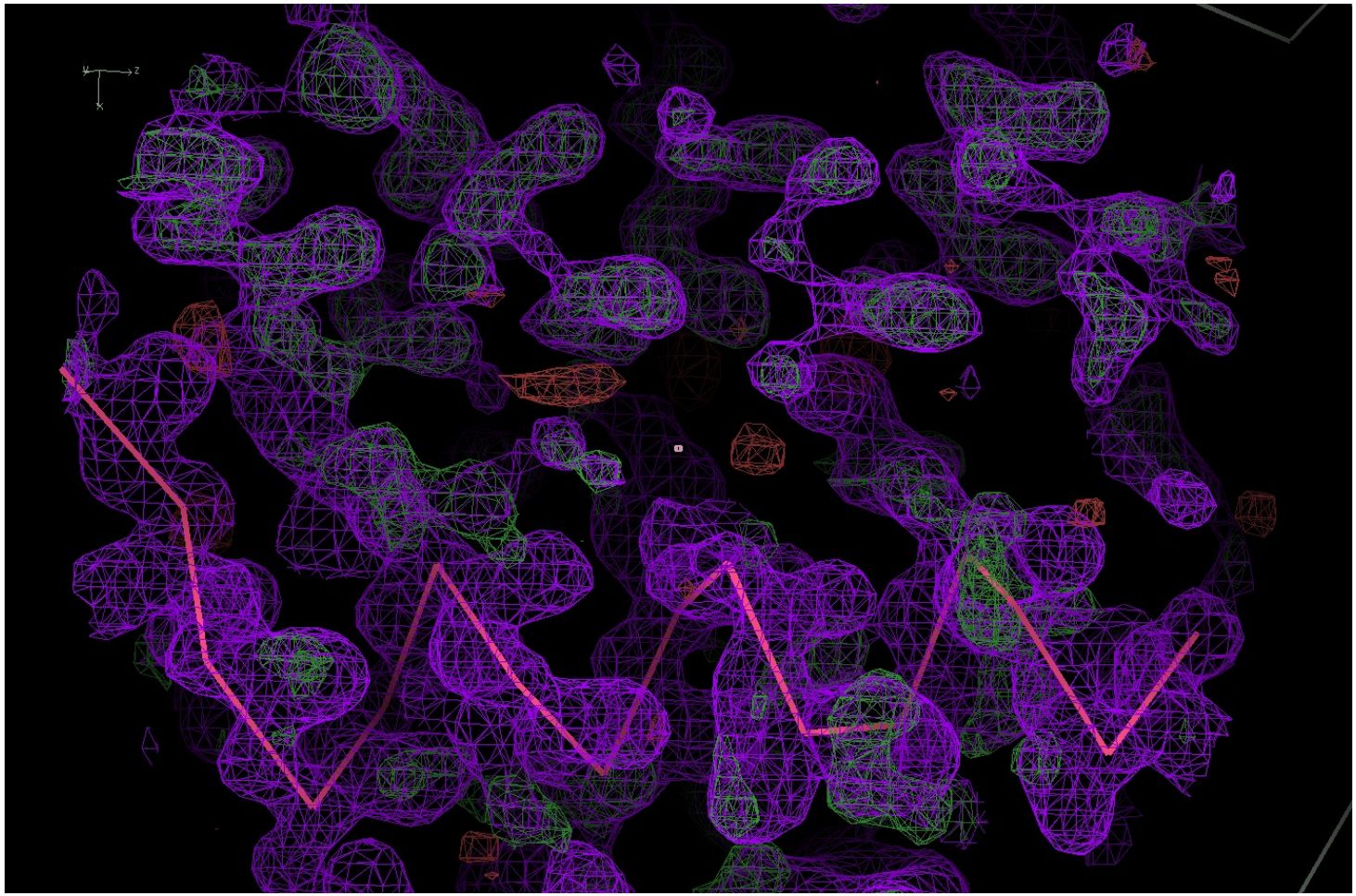
dials.rs_mapper + coot

An example from **Rafael Ciges**, Biomedical Institute of Valencia

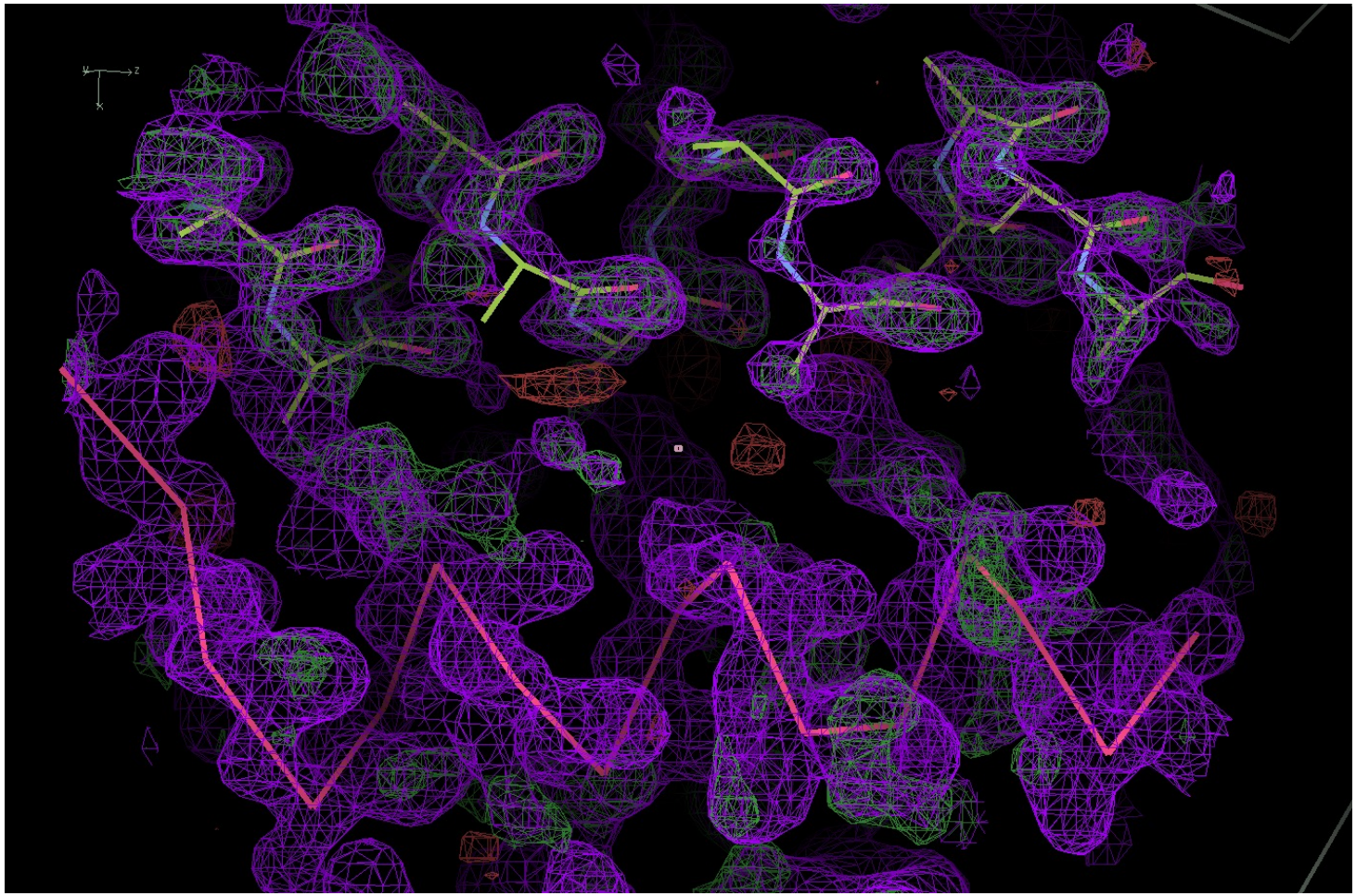


- Space group $P2_12_12$
- Resolution 1.2Å
- Structure was solved with MR
- Preliminary refinement R_{free} = 0.35
- Extra residues were expected compared to MR model

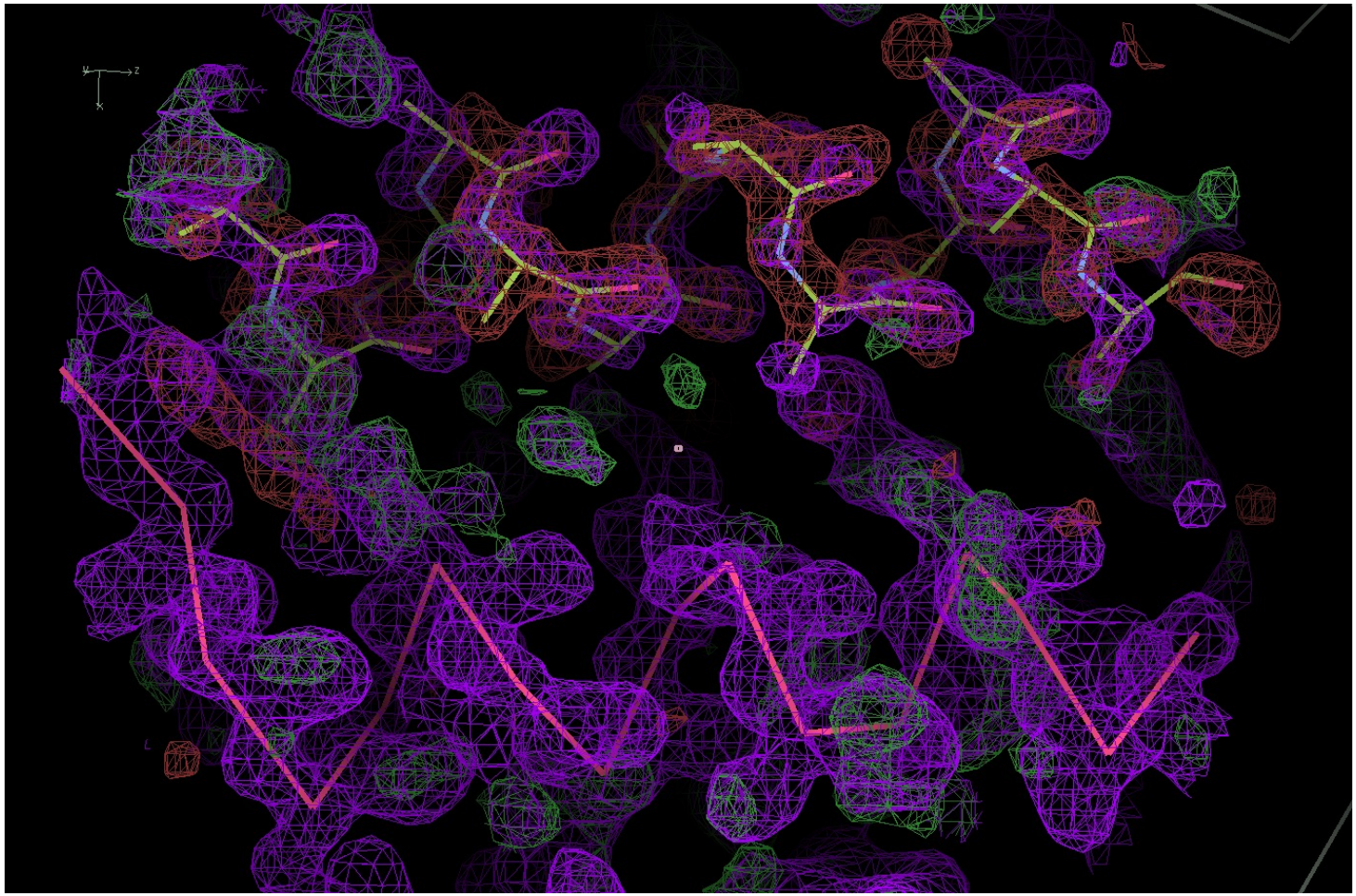
Example 1: after initial refinement



Example 1: helix added



Example 1: after refinement with extra helix

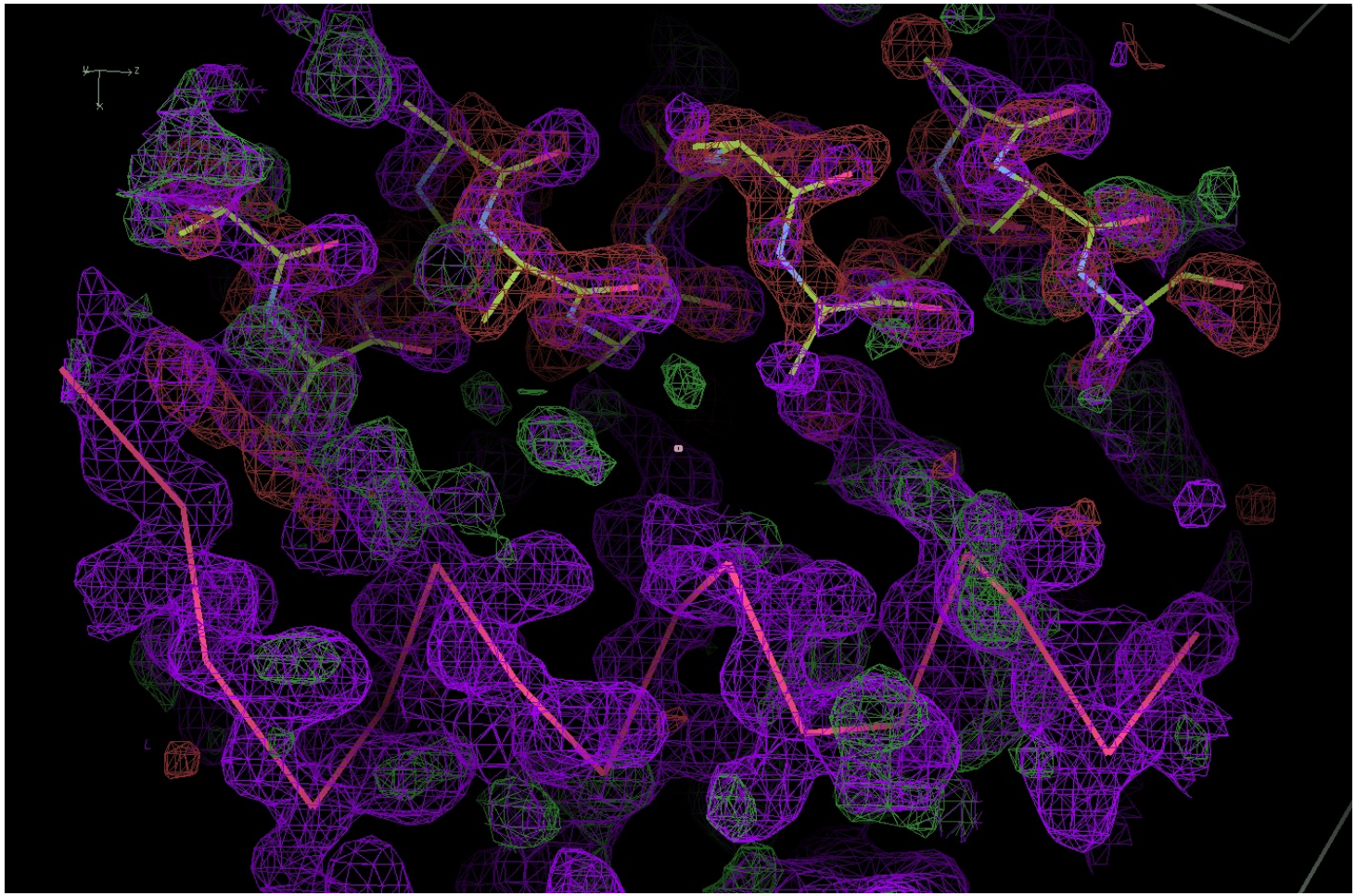


Example 1: demodulation of intensities

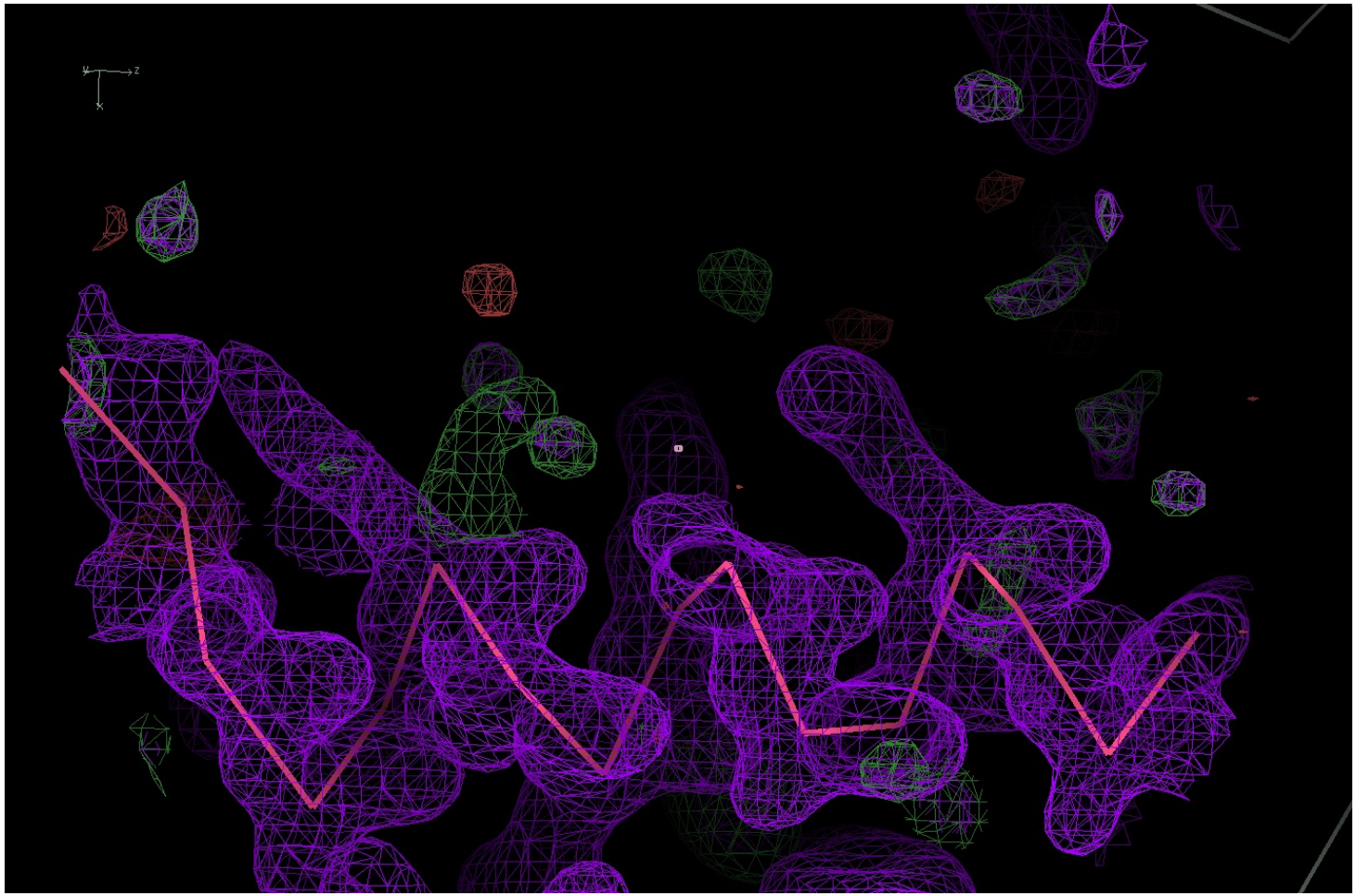
- Data were demodulated to approximately account for disorder

	R	R-free
Original data	0.33	0.34
Corrected data	0.25	0.26

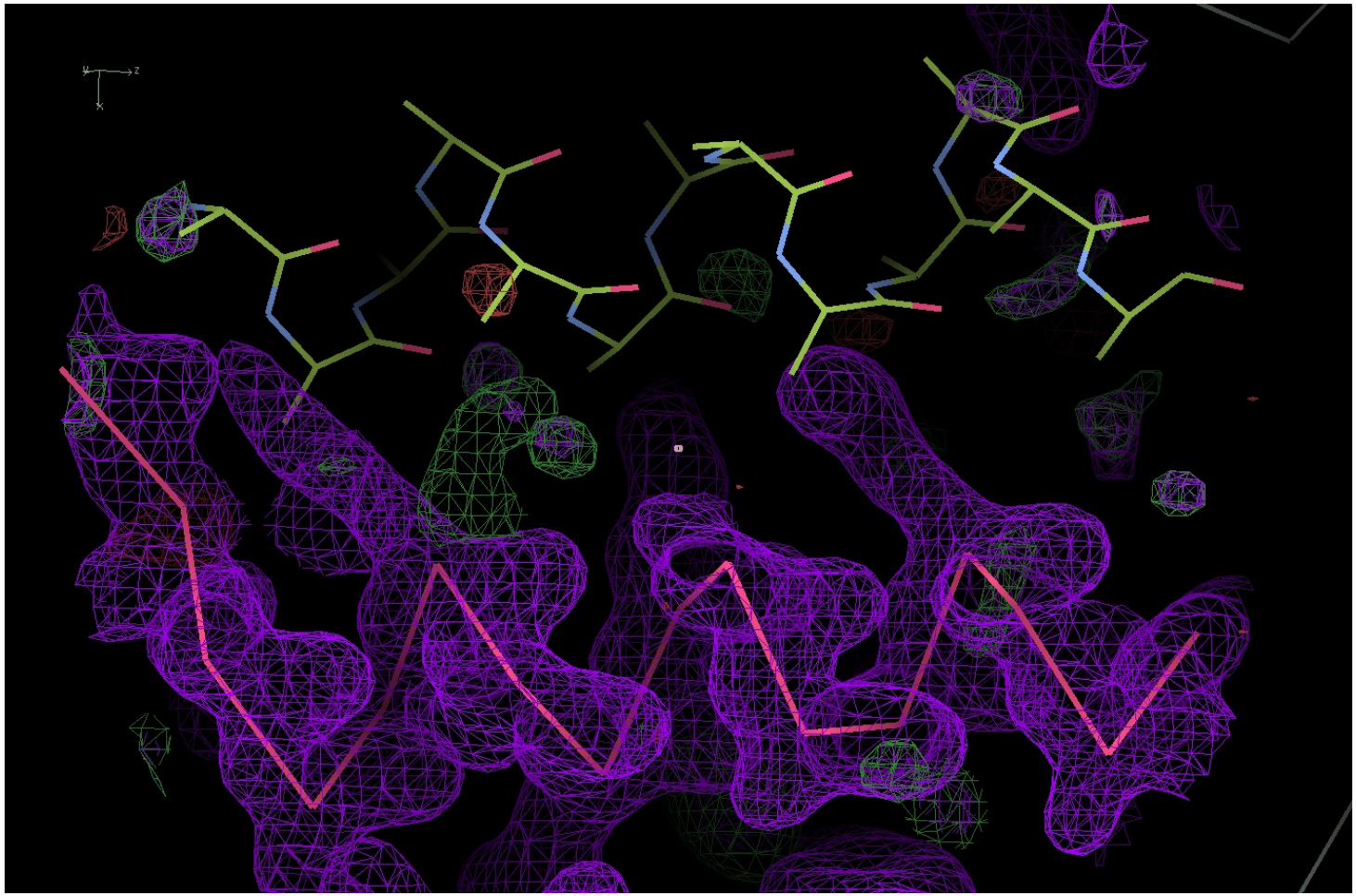
Example 1: after refinement with extra helix



Example 1: after refinement against demodulated data ...



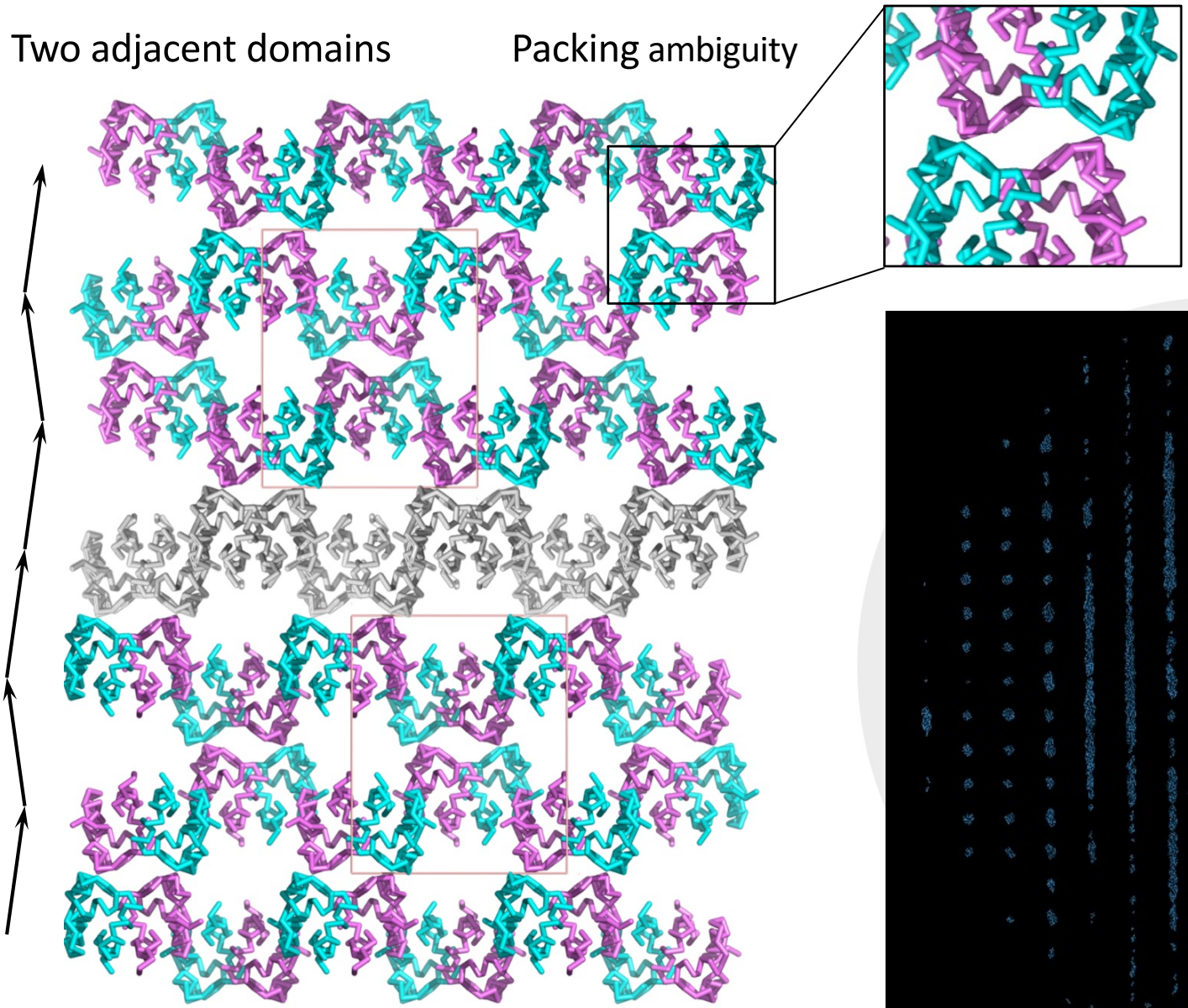
Example 1: ... there is no ED for the extra helix



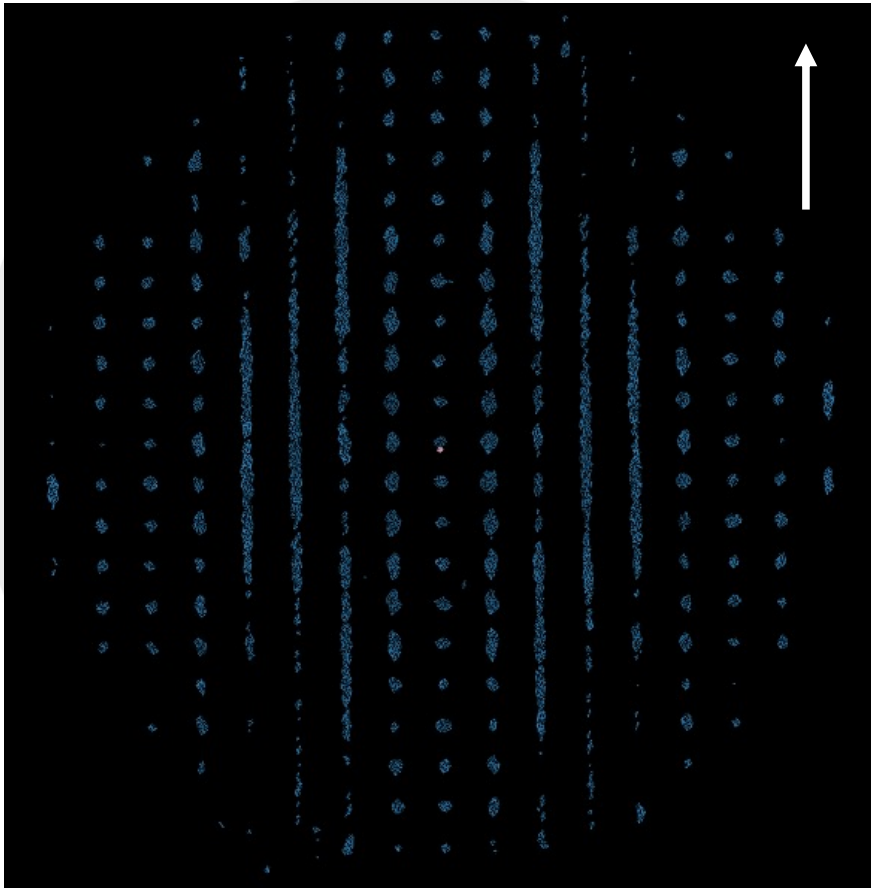
Example 1: ghost density

Two adjacent domains

Packing ambiguity



White arrow - direction in which global periodicity is missing



Example 1: Summary

- Partial disorder in OD structures results in a ghost density
- Structure can be solved and refined ignoring partial disorder
- Demodulation procedure may remove ghost density and therefore help with interpretation of the ED maps
  - » Not always badly needed
  - » Not always works
  - » There are several bespoke scripts around
- Atomic model usually indicate reasons for partial disorder

Example 2: auto-indexing failure

Fast DP @ DIAMOND
(be cautious it is fast fist of all)

Refinement

R_{meas} = 0.12 ?

CC(1/2) = 0.3 at 1.56 Å

R_{cryst} = 0.33 ?

R_{free} = 0.36

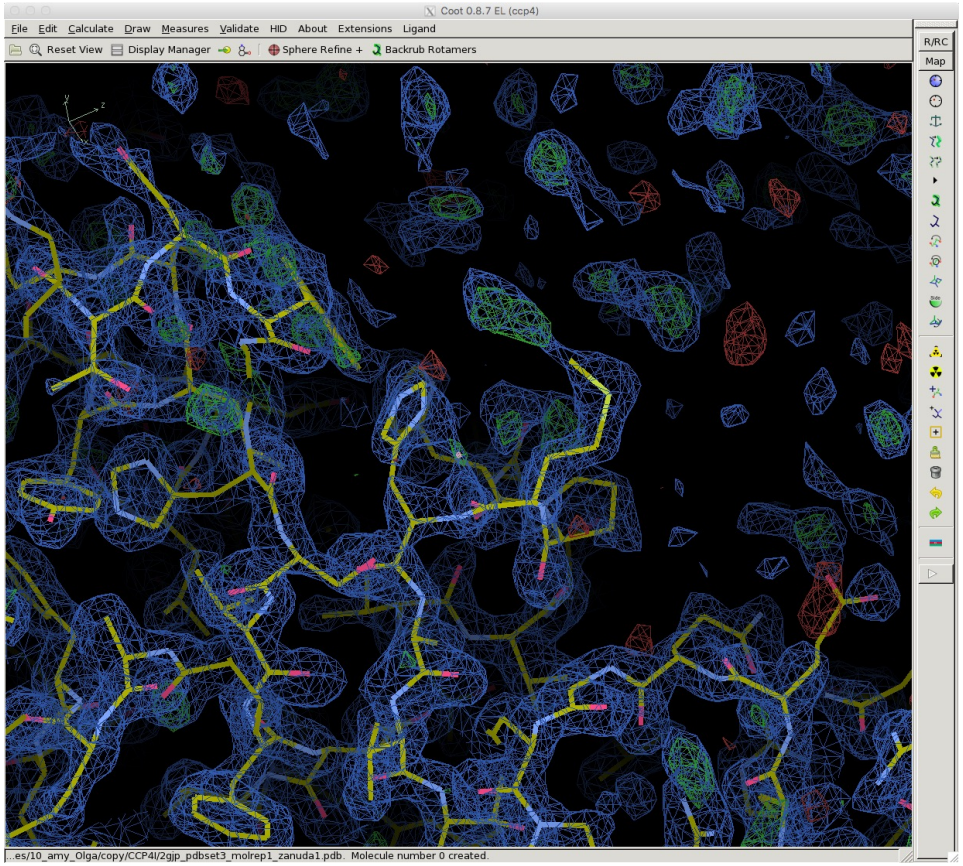
Molecular Replacement ✓

--- Peaks of Rotation Function ---

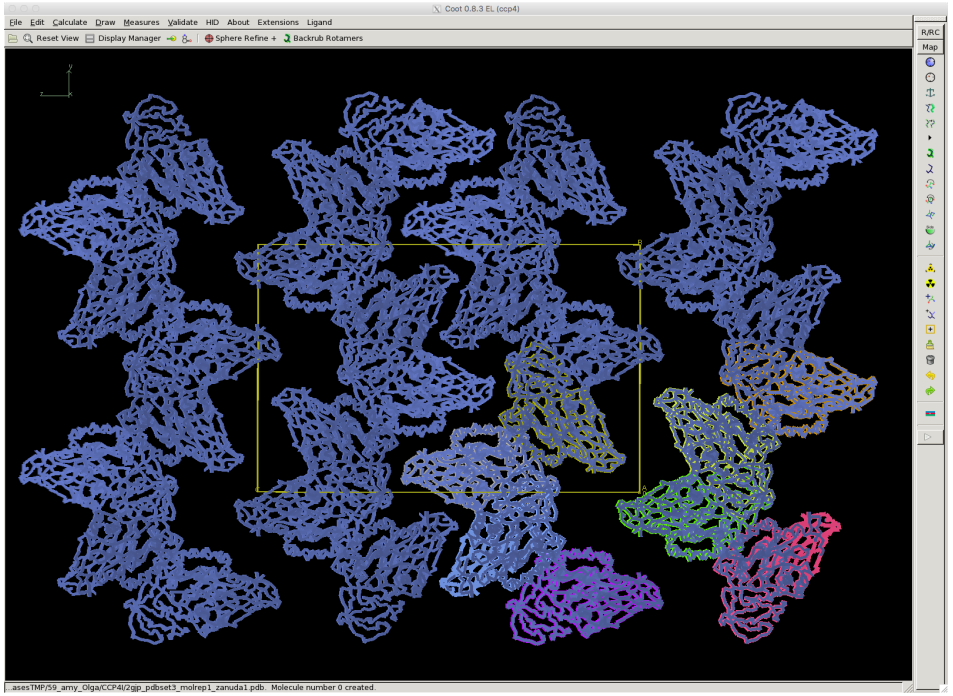
	theta	phi	chi	Rf/sigma
1	63.62	174.24	148.98	13.70
2	80.19	-58.05	61.61	13.63
3	149.48	-148.30	170.26	13.34
4	107.22	84.22	129.22	13.04
5	87.46	75.99	136.16	12.18
6	111.97	-14.20	175.28	12.10
7	157.20	173.73	153.99	11.25
8	58.77	-96.16	51.96	11.24
9	75.76	-63.11	54.46	6.21
10	102.46	82.67	133.90	5.83

Example 2: evidences of wrong indexing

Maps ?



Crystal Packing X

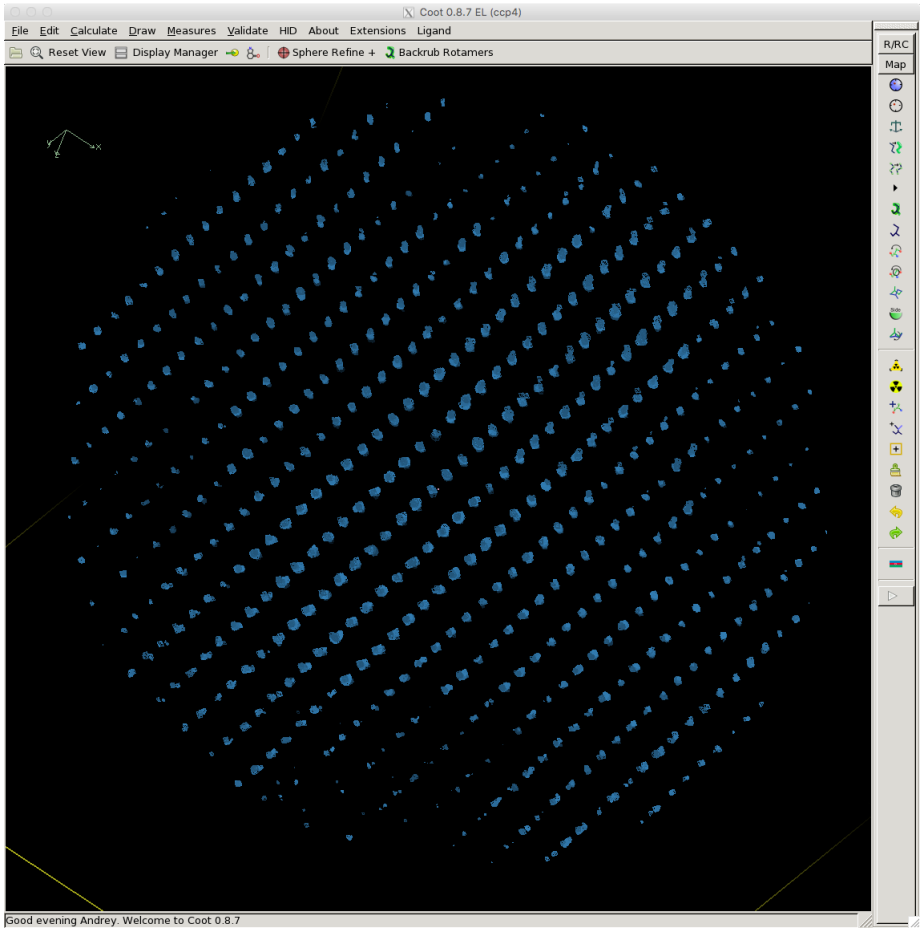


Wrong indexing?

Example 2: evidences of partial disorder

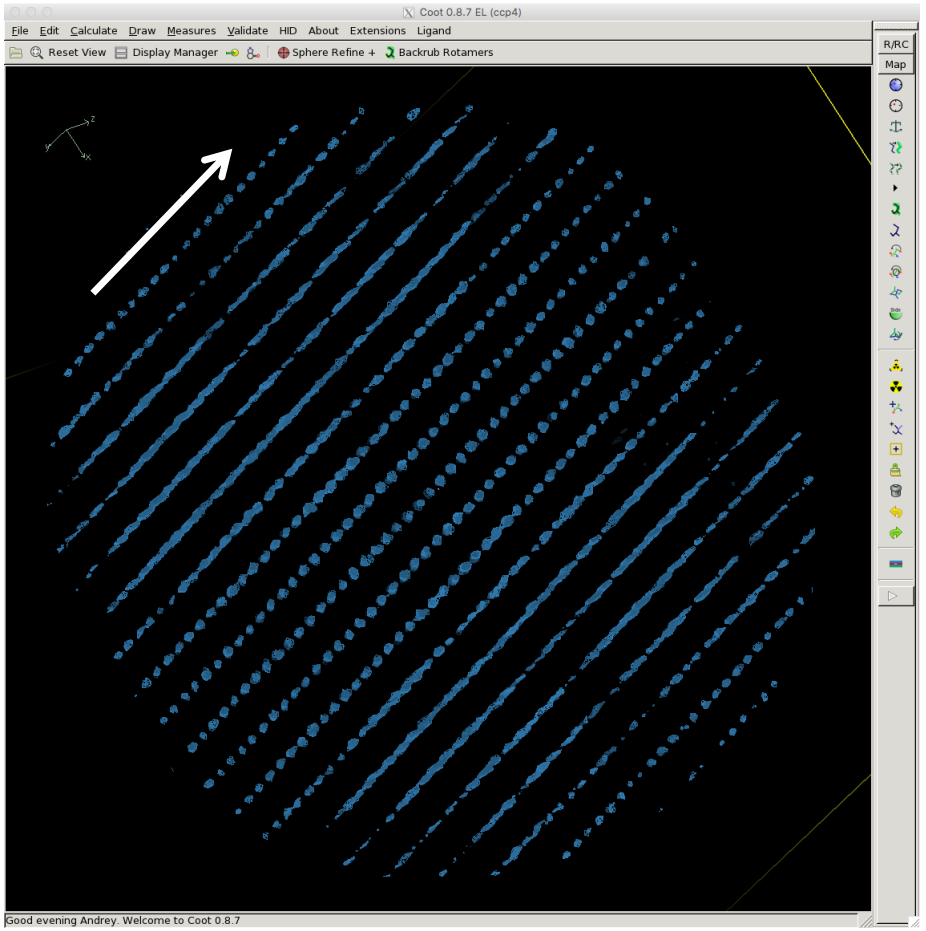
front view

There is global 2D translational symmetry in the plane of figure

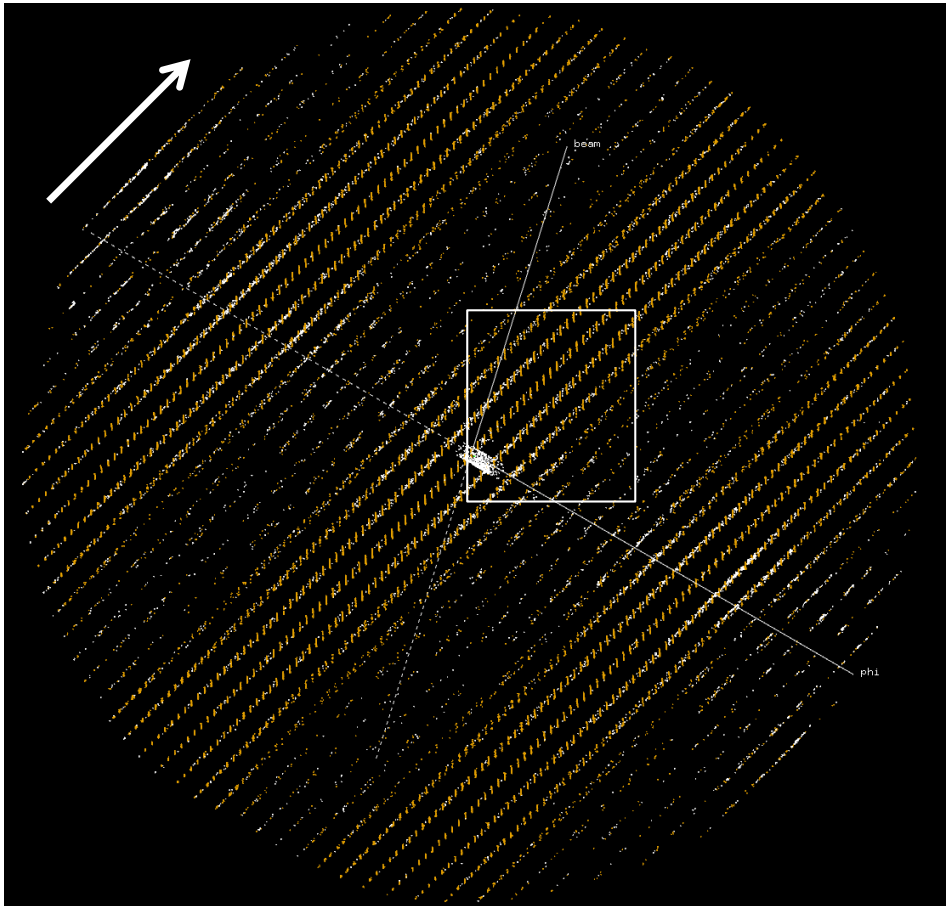


side view

White arrow indicates direction in which translational symmetry is not global (only within individual domains)



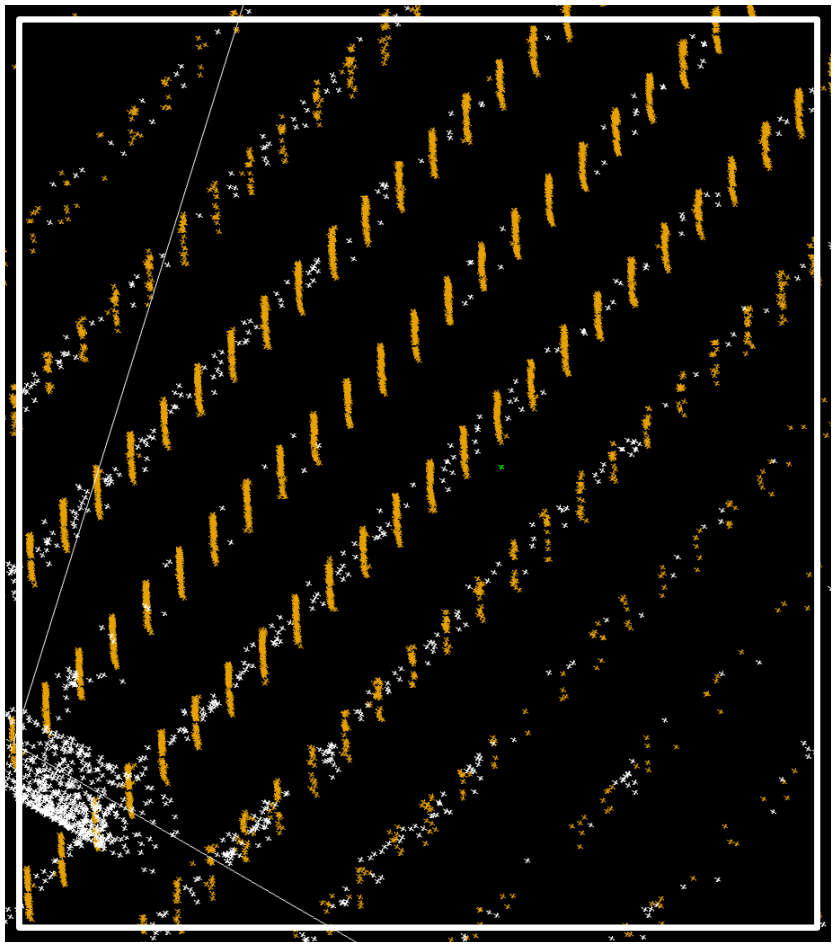
Example 2: correct indexing



White arrow indicates direction in which translational symmetry is not global (only within individual domains)
There are also areas with less spots

White “spots” are not indexed; actually, these are tails of diffuse reflections

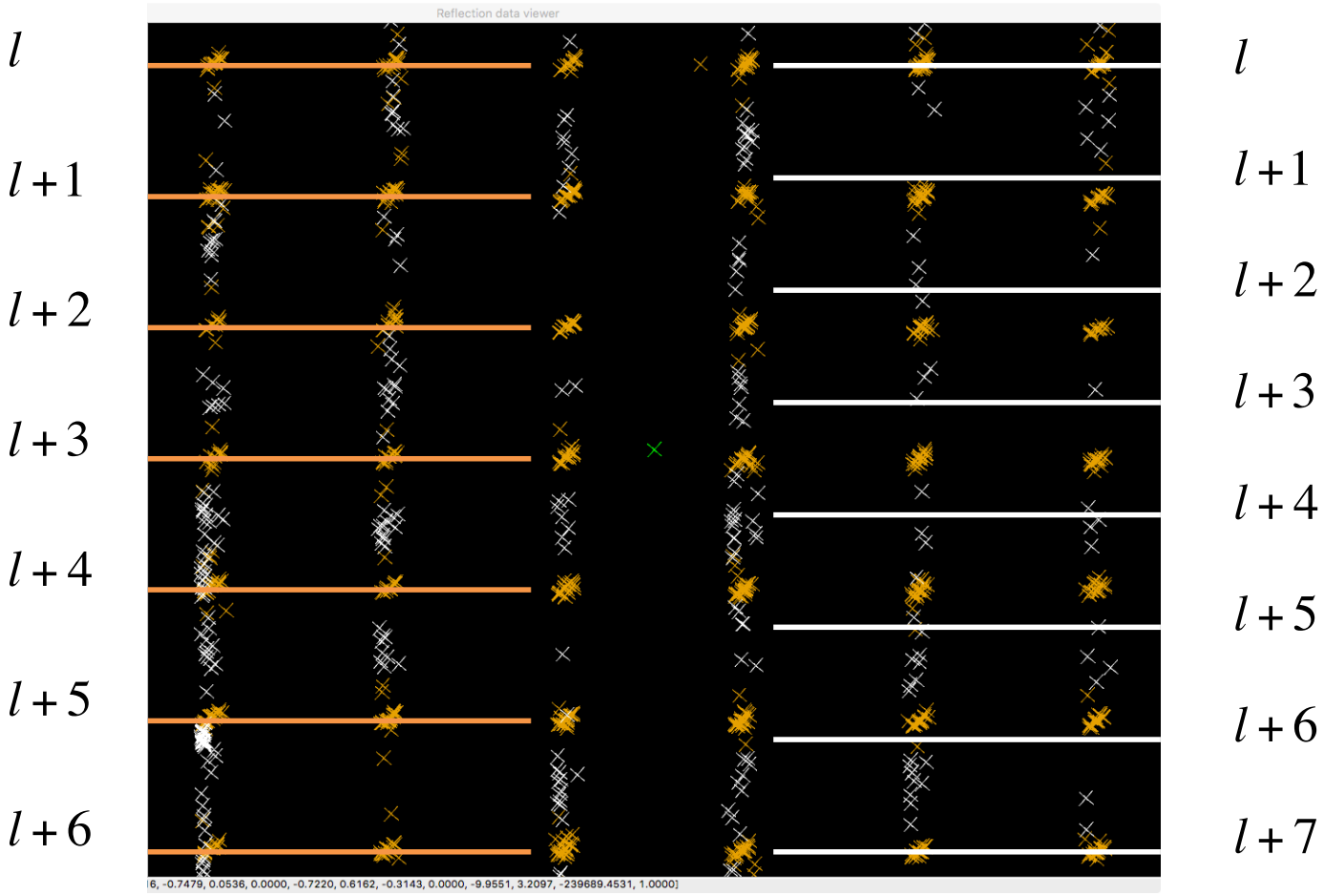
Indexing program may take them for real spots and produce wrong result.



Example 2: what initially was wrong

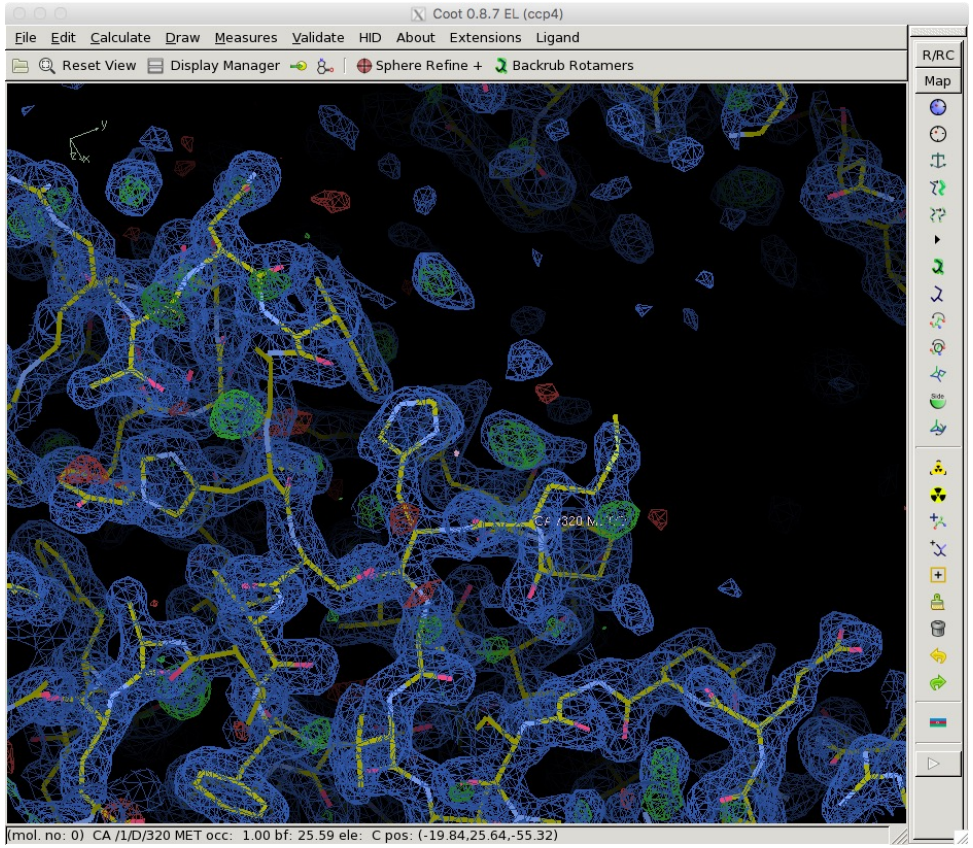
correct indexing

wrong indexing

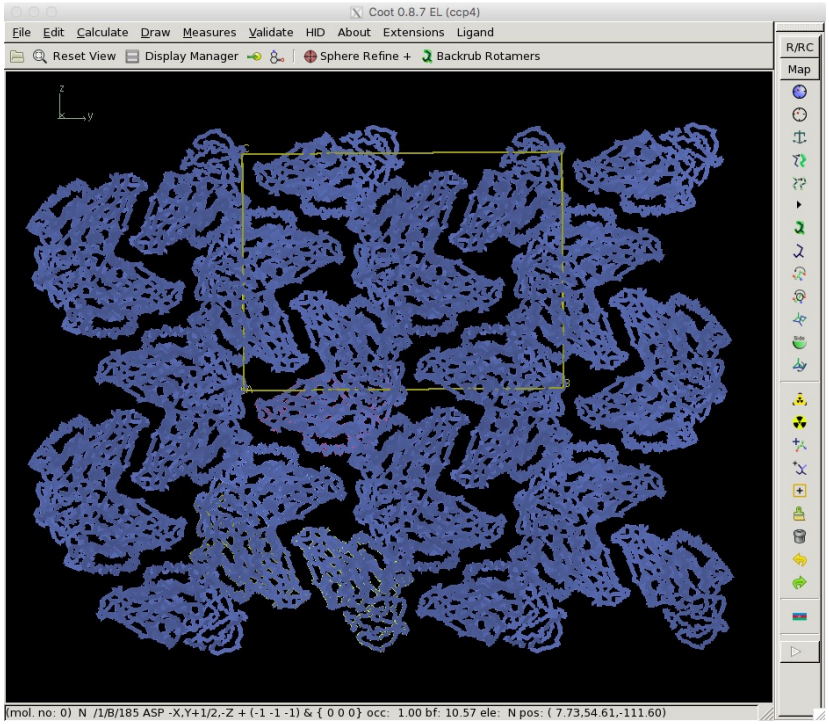


Example 2: happy end

Maps ✓



Crystal Packing ✓

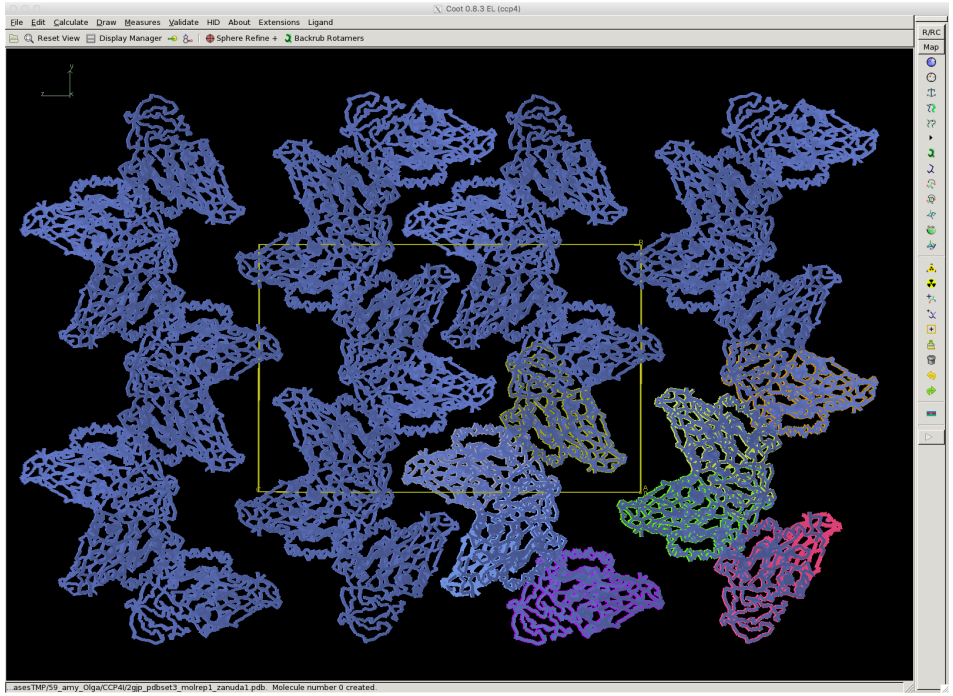


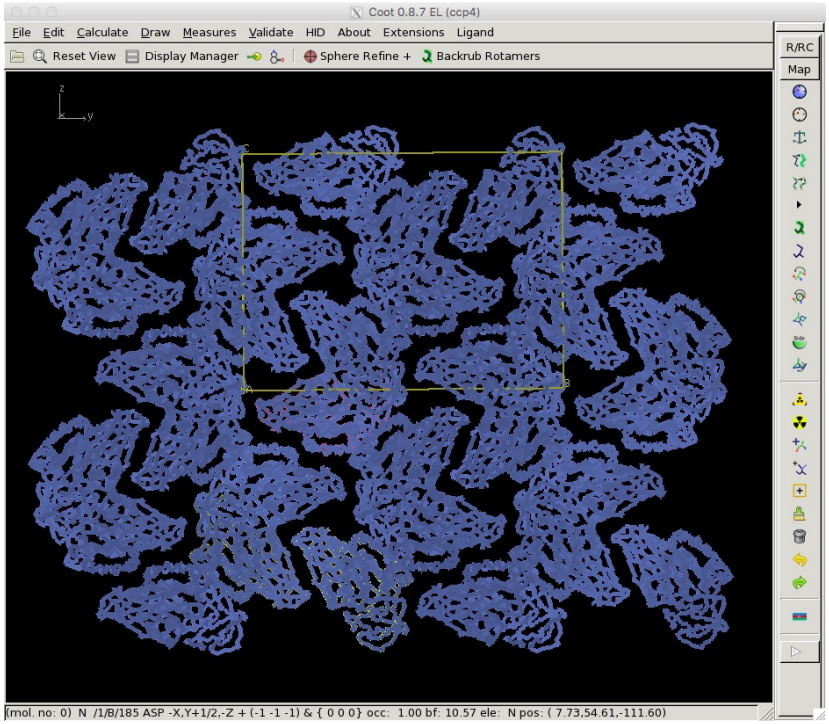
Refinement ✓

R_{cryst} = 0.23

R_{free} = 0.26

Example 2: wrong and correct





Example 2: Summary

- Partial disorder a frequent reason of indexing failure
- Use 3d viewers for diagnostics
- Warning: high contrast in MR can be obtained even for wrongly indexed data provided that the search model is highly similar to the target
- Look at crystal packing and solvent area

Partial disordered OD structures

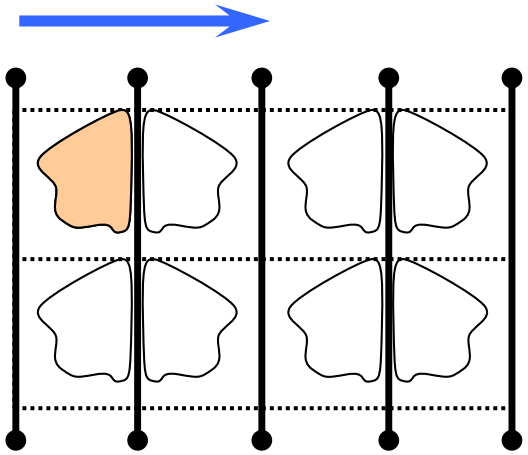
- Signs: (some) rows of reflections have tails or merge together
- Data processing
  - Indexing can go wrong (use higher "gain" parameter, merge several adjacent images together etc. to get it right)
- Structure solution:
  - Molecular Replacement - yes
  - Experimental phasing – might be problematic
- Refinement / model building:
  - Some features of electron density may be non-interpreted (ghost density)
  - Expect higher R-factors
- Crystals with translocation defects
  - Term usually used in MX for partially disordered OD-structures of type I

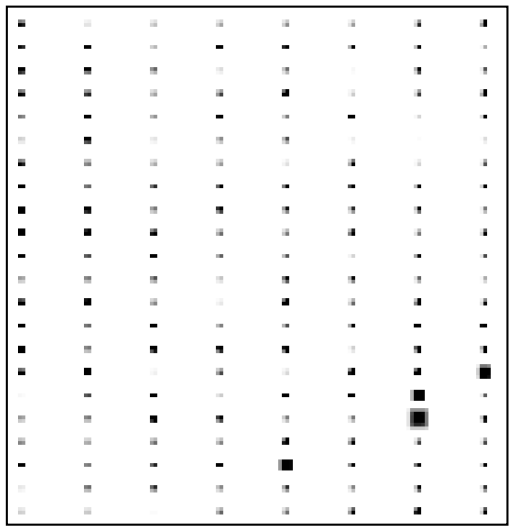
Pseudo-translation

- Presence of weak reflections
- Effect on indexing

Pseudotranslation

Crystallographic translation





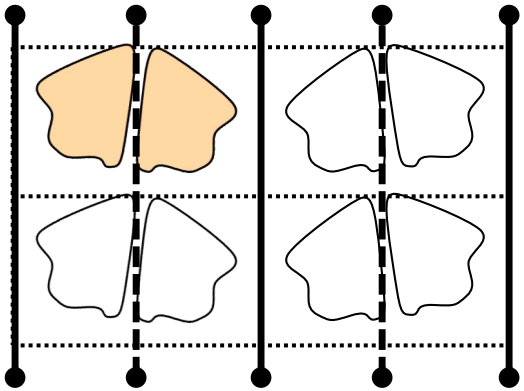
No pseudotranslation

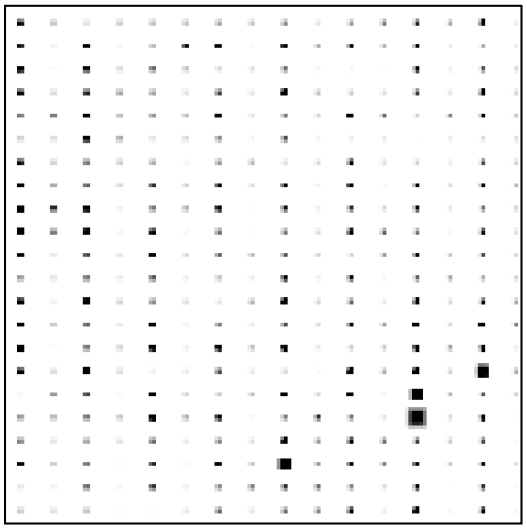
c

*c**

Crystallographic translation

Pseudo-translation





Pseudotranslation

c’ = 2 *c*

c’* = *c**/ 2

Planes 2L+1 contain weak reflections

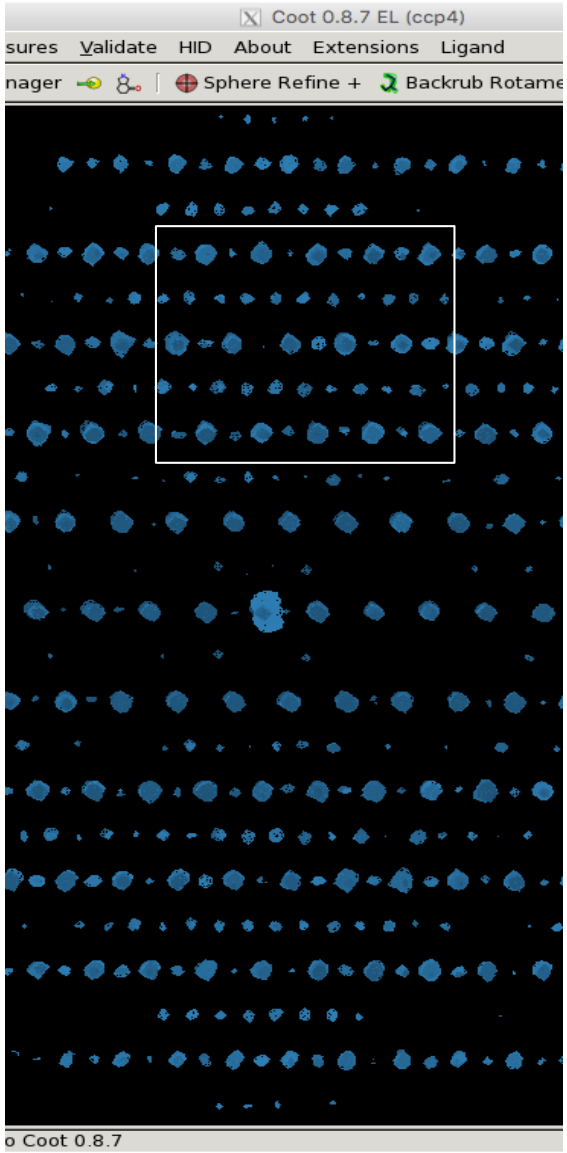
Example: two pseudo-translation vectors

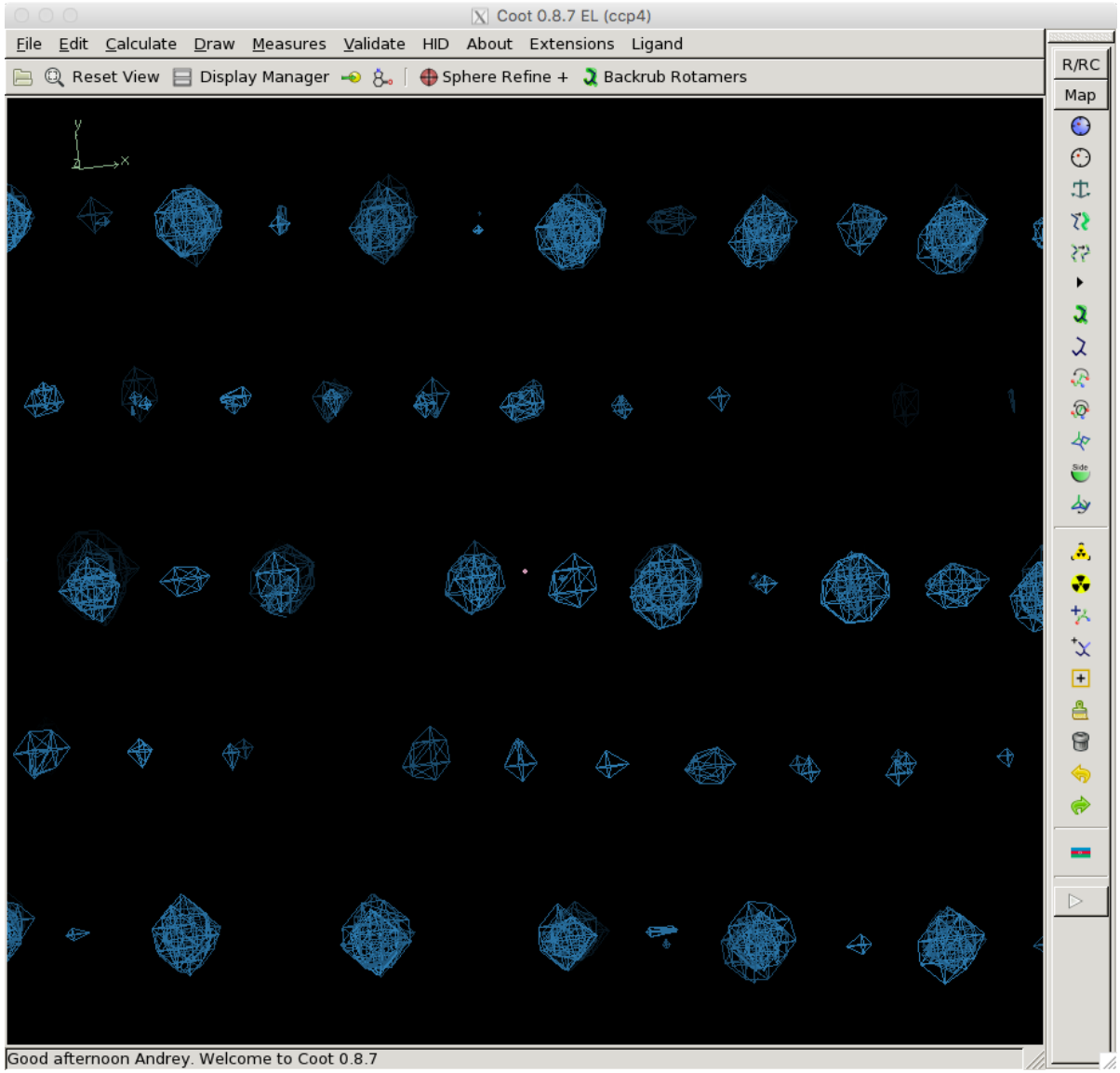
	space group	*a* (Å)	*b* (Å)	*c* (Å)
All reflections	*C* 2 2 2	74.9	122.8	125.0
Strong reflections only	*I* 2 2 2	37.5	61.4	125.0

Example from Victor Lamzin, YSBL-DESY

» structure solved using SAD

Example: two pseudo-translation vectors

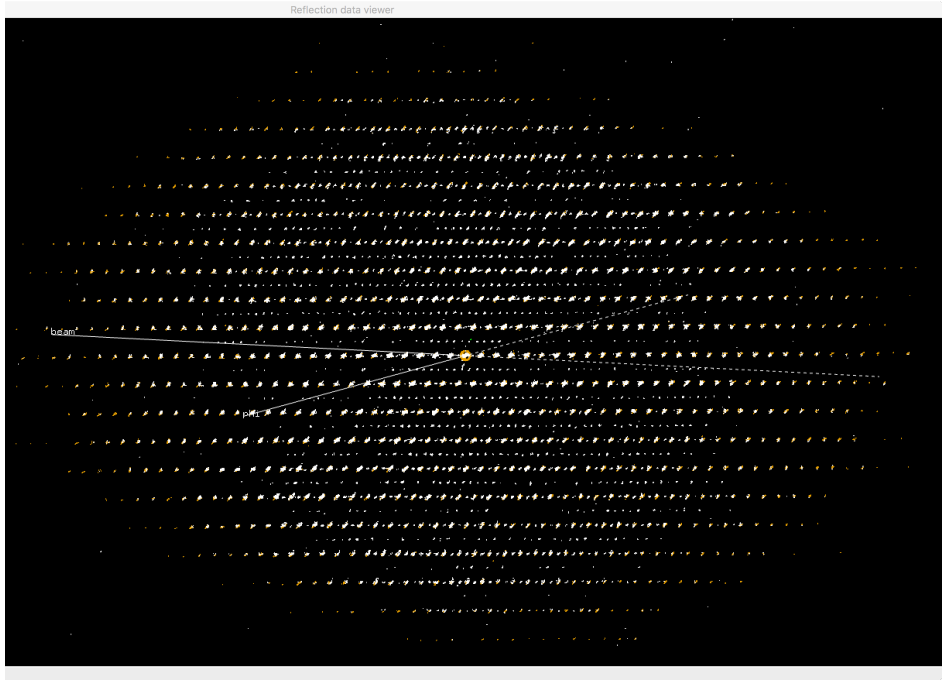


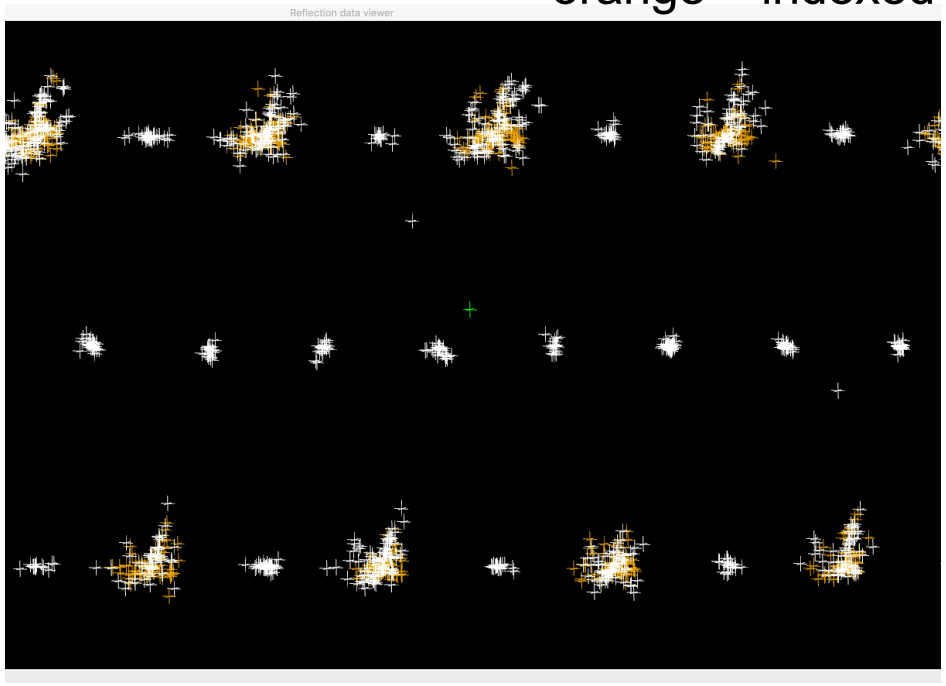


Example: two pseudo-translation vectors

Images imported as they were, oscillation 0.1°

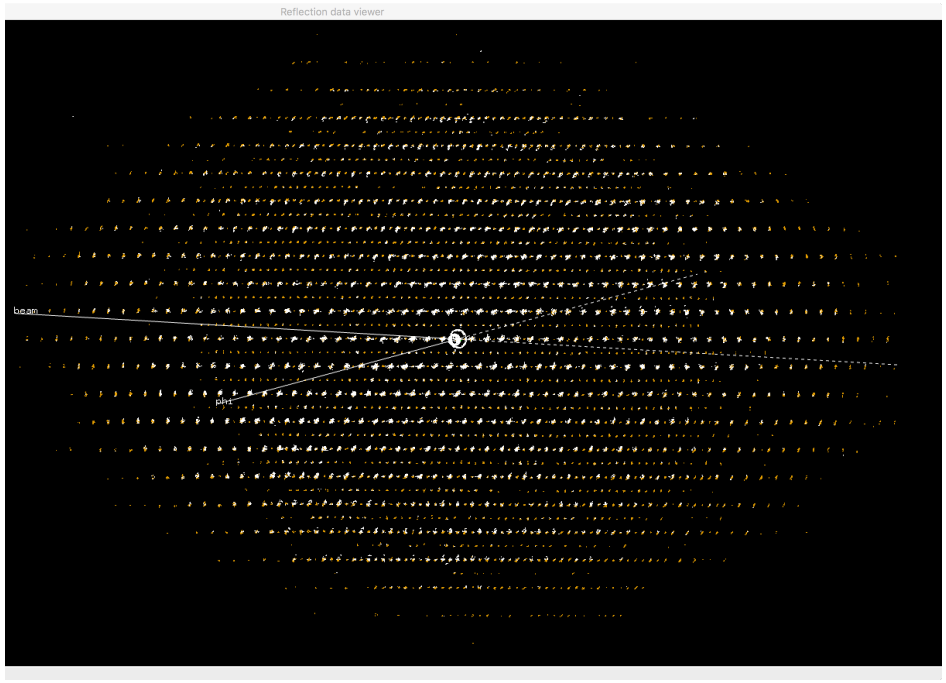
white – not indexed
orange – indexed

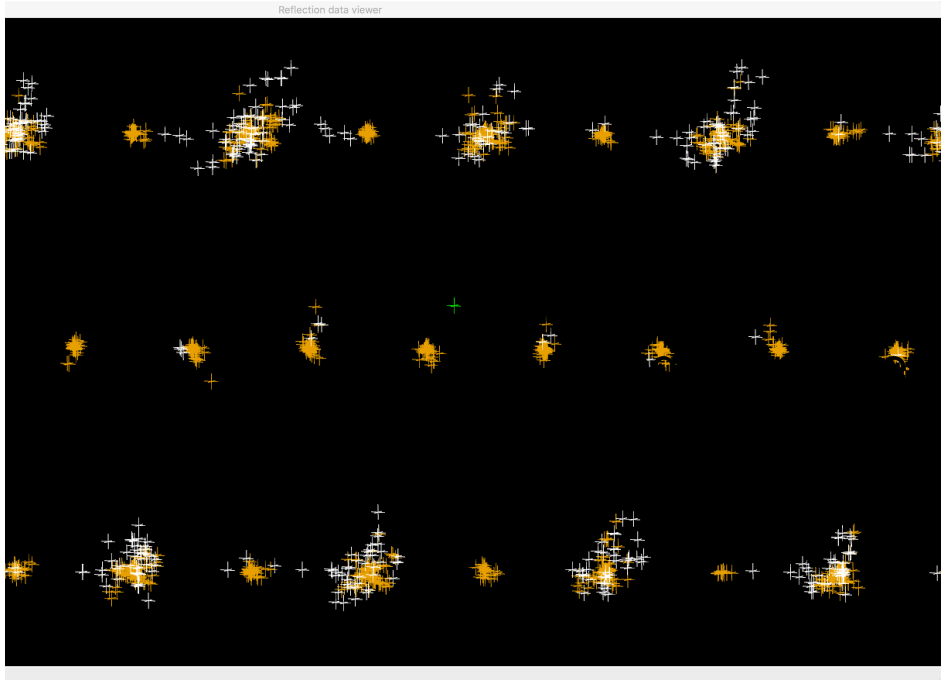




Example: two pseudo-translation vectors

Merged each 5 adjacent images to make oscillation 0.5° , then imported





Pseudo-translation

Week reflections may confuse indexing programs

Visual control using 3D viewers is useful

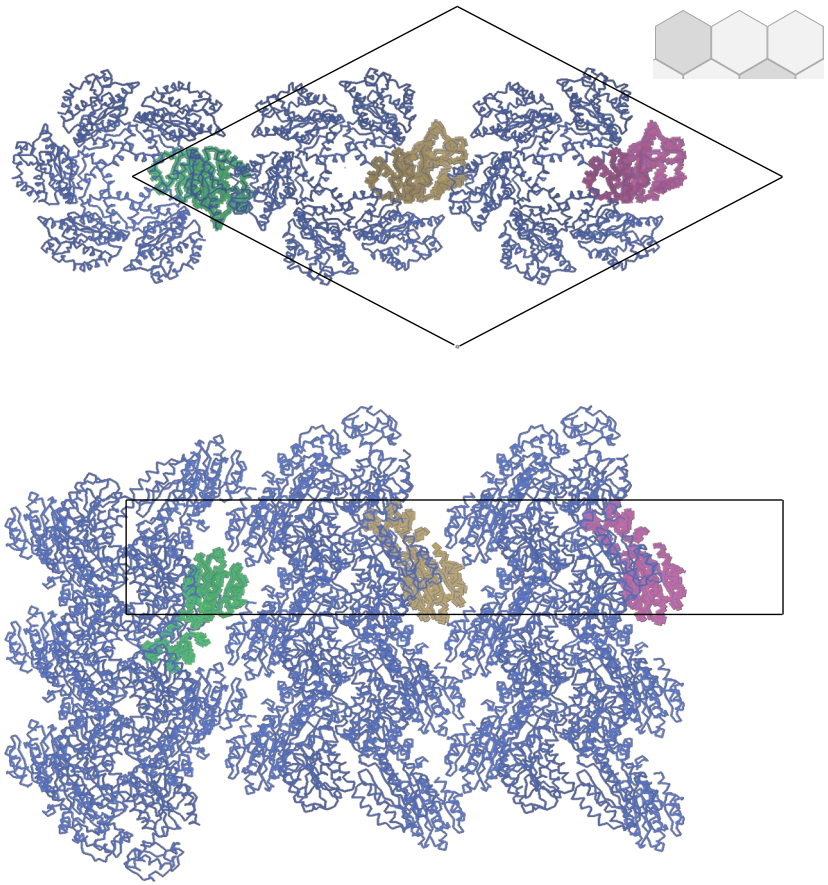
- check if pseudo-translation is not overlooked
- check if pseudo-translation is not an indexing artefact

How important is it to use weak reflections?

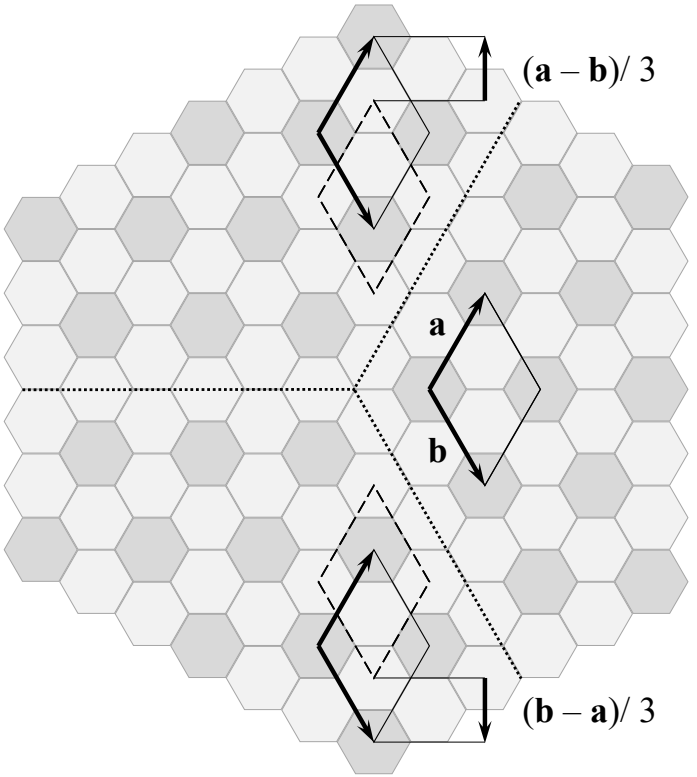
- usually improve both density and refinement stats
- sometimes ignored to simplify the first steps of structure solution and used later
- there are examples when these only make refinement stats worse
  - » the last example - "fake pseudo-translation" - suggests a possible reason

Fake pseudotranslation

An example from
Maria Jose Pedroza Romo and **James Moody**
Brigham Young University, Utah

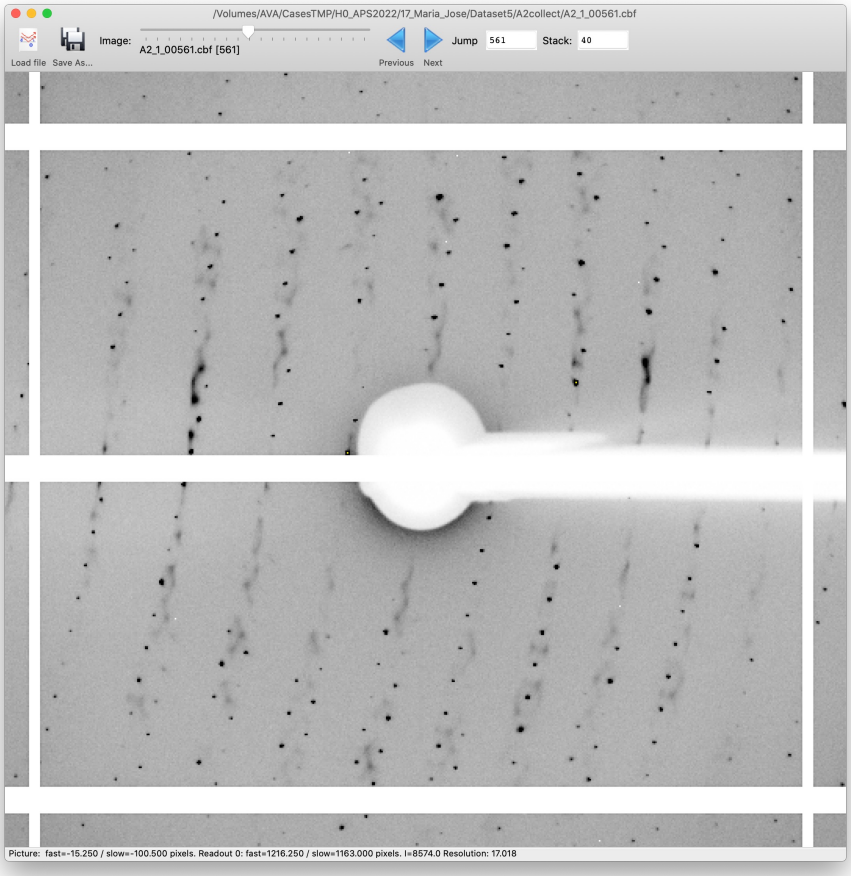


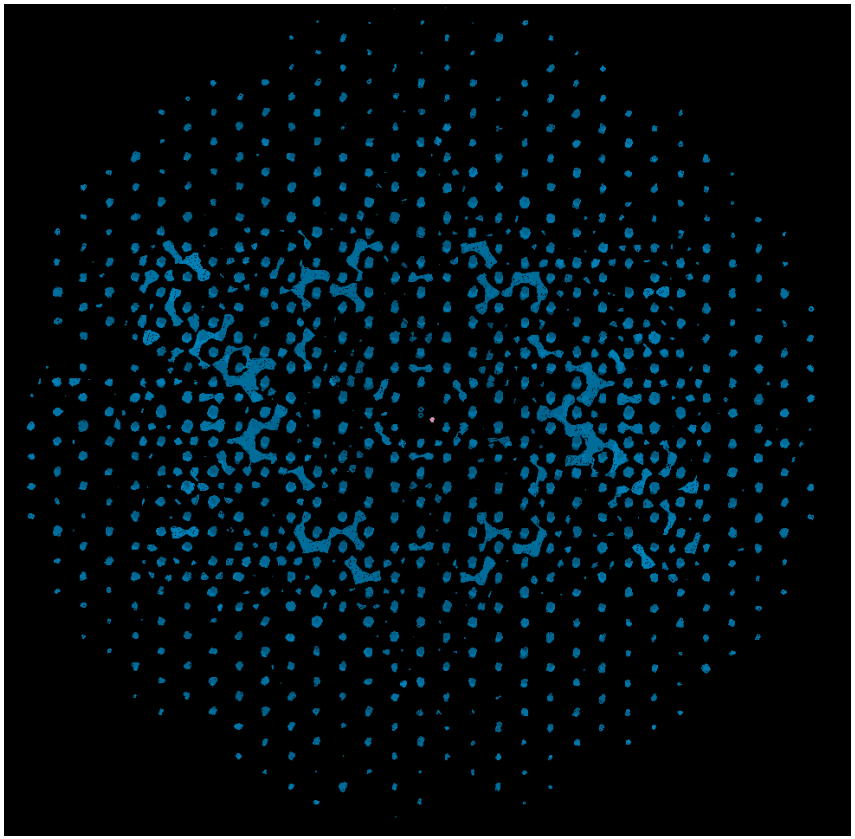
"Pseudotranslation" at the level of crystal domains



"strong" reflections :	$h - k = 3n$	$I_{\text{obs}} = I_{\text{dom}}$
"weak" reflections :	$h - k = 3n \pm 1$	$I_{\text{obs}} = \frac{V_{\text{dom}}}{V_{\text{coh}}} I_{\text{dom}}$

4. Fake pseudotranslation





MR and initial refinement: “strong” reflections (completeness 33%), R=0.26, R_{free}=0.28

Rescaling: “weak” reflections to Fcalc

Final refinement: “strong” + “weak” (completeness 100%), R=0.27, R_{free}=0.31

END

You are welcome to discuss your problematic data